

In re Application of:  
Fenical et al.  
Application No.: 09/991,518  
Filed: November 16, 2001  
Page 7

PATENT  
Attorney Docket No.: UCSD1630-1

BEST AVAILABLE COPY

### REMARKS

These remarks are in response to the Office Action mailed February 15, 2005. Claims 1 and 18 have been amended. Claims 25-32 have been added. Claims 13-16 are canceled without prejudice. Claims 2-12, 19, and 21-22 were previously canceled. No claims are withdrawn. Thus, claims 1, 17-18, 20 and 25-32 are presently pending. Hence, Applicants submit that a new search is not necessary, and that all amendments are supported by the previously pending claims and originally filed specification.

Also, SEQ ID NOs: 3 (accession no. AY040619), 4 (accession no. AY040620) and 5 (accession no. AY040622) were first seen on the NCBI database on August 23, 2002. The deposit satisfies the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure (1977). Applicants amended the specification (paragraphs [0026] to [0028]) in the response to the Office Action mailed September 15, 2003 (page 2-3 and Exhibit A of the response) to reflect the above.

## **II. Amendments to the Specification**

Paragraph [0026] has been amended to particularly point out the subject matter of the claimed invention, which are 16S rRNA sequences from a strain of *Salinospora* containing certain signature nucleotides. The 16S rRNA sequences are patentable irrespective of how the signature sequences are identified. Applicants note, that in order to identify and characterize a microbial organism, it is standard in the art to perform a computational alignment analysis against other known 16S rRNA gene sequences, including but not limited to, the *E.coli* 16S rRNA sequence. The *E.coli* 16S rRNA sequence is, however, the reference sequence with which all other 16S rRNA sequences are compared. This aspect will be made more clear in the following response and in the expert Declaration under 37 CFR § 1.132 of Dr. Stephen Giovannoni (see Appendix and Exhibits A, B and C). Therefore, paragraph [0026] has been amended to particularly reference *E.coli* as the standard for the 16S rRNA gene sequences.

Table 3 of the specification (page 12) has been amended to include another column indicating the actual nucleotide positions of the signature nucleotides of the claimed *Salinospora* actinomycete isolates. Applicants have amended Table 3 to eliminate the confusion of the use of the phrase “*E.coli* numbering system” with respect to the claimed subject matter. The signature sequences of the *Salinospora* isolates of the instant invention are not dependent on the *E.coli* 16S rRNA sequence, the *E.coli* sequence is only used as a relative marker with respect to the *Salinospora* sequences. That is, the signature sequences of the instant invention are “signature” because they do not exist in *Micromonosporaceae* 16S rRNAs sequences, and not because they are identified based on the *E.coli* 16S rRNA sequence. Thus, the amendment to Table 3 does not add new matter.

## II. Amendments to the Claims

Claims 1 and 18 have been amended. Claims 1 and 18 have been amended to improve their form. The claimed invention relates to 16S rRNA signature nucleotides of at least 3 *Salinospora* isolates, e.g. SEQ ID NOs: 3, 4 and 5.

Claim 1 has been amended and now recites, “[a]n isolated marine actinomycete having an obligate requirement of sodium for growth, wherein the marine actinomycete is a strain of *Salinospora* comprising 16S rRNAs SEQ ID NO:3, 4 or 5.” Amended claim 1 contains patentable subject matter, irrespective of the *E.coli* 16S rRNA sequence numbering system, because as discussed above, the signature sequences are “signature” because they are not found in *Micromonosporaceae* genera, and not due to their numbering system.

The amendments to claims 1 and 18 are fully supported by the specification, e.g. paragraphs [0026] and [0027]. Paragraphs [0026] and [0027] were first amended in the response mailed December 15, 2003 to the Office Action mailed September 15, 2003, to include the sequence identifiers (SEQ ID NOs: 3, 4 and 5) of the described isolates (accession numbers: AY040619, AY040620 and AY040622, respectively). The accession numbers were disclosed in the application as filed. Hence, it is clear that Applicants were in possession of the nucleotide sequences at the time the application was filed.

Claims 13-16 have been canceled because they have been incorporated into independent claim 1. Claims 25-30 have been added.

Claims 25-30 do not add new subject matter because they simply claim that which was previously presented in pending claim 1.

Accordingly, Applicants submit that a new search is not necessary, and that all amendments to the claims are supported by the application as filed and by the previously pending claims.

### **III. Rejection Under 35 U.S.C. §112**

#### **A. 35 U.S.C. §112, first paragraph (written description)**

Claims 1, 13 and 17-18 and 20 stand rejected under 35 U.S.C. §112, first paragraph as allegedly failing to comply with the written description requirement. Applicants respectfully traverse this rejection.

Claims 1 and 18 have been amended and claims 25-30 have been added. Claims 1 and 18 have been amended per the suggestion of the Office Action on page 4.

Claim 13 has been canceled, making the rejection with regards to claim 13 moot.

Claim 1 has been amended to incorporate the limitation that the isolates contain SEQ ID NOs: 3, 4 or 5. Claim 1 has also been amended to delete the recitation relating to identification of the signature sequences based on the *E.coli* 16S rRNA numbering system in the Ribosomal Database Project (RDP). Attached herein is a Declaration under 37 CFR § 1.132 by Dr. Stephen Giovannoni (see Appendix and Exhibits A-C).

Dr. Giovannoni is an expert in the area of microbial identification and characterization. Attention is drawn to paragraphs 3, 4 and 5 of the Declaration of Dr. Giovannoni, whereby Dr. Giovannoni states:

3. First, as "one skilled in the art," I declare that that the universal "gold standard" for microbial classification is a computational alignment analysis against existing 16S rRNA gene sequences. This same "gold standard" was utilized in the above-identified patent application to identify and characterize the signature sequences in the *Salinospora* isolates.
4. Second, Bergey's Manual of Systematic Bacteriology (George M. Garrity, David R. Boone, Editors) is a text used worldwide, and supports the above mentioned "gold standard" (i.e. microbial identification and classification schemes rely on 16S rRNA sequences). Bergey's Manual of Systematic Bacteriology, is also the taxonomic hierarchy used in the Ribosomal Database Project (RDP). The RDP provides aligned and



annotated rRNA gene sequences, and is made available through the RDP website (<http://rdp.cme.msu.edu/>). See also Cole et al, "The ribosomal database project (RDP-II): sequences and tools for high-throughput rRNA analysis," *Nucleic Acids Research*, 33:D294-296 (2005), which is hereby attached as Exhibit B. The signature nucleotides of the *Salinospora* 16S rRNA sequences claimed in the above identified application were aligned against other 16S rRNA gene sequences, including members of Micromonosporaceae in the RDP (see paragraph [0028] and Table 3 of the specification).

5. Third, Brosius et al. (1978) was first to publish the complete 16S rRNA of *E.coli* (please see Brosius et al., "Complete nucleotide sequence of 16S ribosomal RNA gene from *Escherichia coli*," *Proc. Natl. Acad. Sci. USA* 75(1):4801-4805, which is hereby attached as Exhibit C). The Brosius 16S rRNA *E.coli* gene sequence is the reference sequence used worldwide and is the sequence found on the RDP. This is true regardless of any changes and updates to the sequence. To put it another way, the *E.coli* 16S rRNA gene sequence on the RDP is "permanent and invariant." Therefore, the Applicants mention of "*E.coli* positions 27-1492" is a reference to the Brosius' *E.coli* 16S rRNA gene sequence. Those skilled in the art will immediately understand that the Applicants were in possession of the claimed subject matter based on a comparison of the invention *Salinospora* 16S rRNA sequences with *E.coli* 16S rRNA sequence.

Thus, as demonstrated, in part, by the Declaration of Dr. Giovannoni above, computational alignments using small rRNA gene sequences are universally accepted, and is the "gold standard;" that Bergey's Manual of Systematic Bacteriology supports the "gold standard" use of 16S rRNA in computational alignment analysis, and that this taxonomic hierarchy is used in the RDP; that the Brosius et al. (1978) *E.coli* 16S rRNA gene sequence is the same *E.coli* 16S rRNA sequence used worldwide today, regardless of any changes that may be found in the *E.coli* sequence after that date; and that the Brosius et al. *E.coli* 16S rRNA sequence is the same sequence referenced in Applicants application and found in the RDP.

Additionally, Carl Woese, an eminent scholar in the area of microbial diversity and winner of the National Medal of Science (2000), states that:

rRNAs are at present the most useful and most used of the molecular chronometers. They show a high degree of functional constancy, which assures relatively good clocklike behavior. They occur in all organisms, and different positions in their sequences change at very different rates, allowing most phylogenetic relationships (including the most distant) to be measured, which makes their range all-encompassing. Carl R. Woese, "Bacterial Evolution," *Microbiological Reviews*, pp. 221-271 (1987); see Attachment 1.

The statement that "different positions in their sequences change at very different rates," is a reference to signature nucleotides found within particular groups or clades of organisms. Stackenbrandt et al. (1997) supports this same notion that "[s]ignature nucleotides are derivatives of the classification process; i.e. signatures are determined for those organisms that are contained within a particular data set." Stackenbrandt et al., "Proposal for a new hierarchic classification system, *Actinobacteria* classis nov.," *Int. J. Systematic Bacteriology*, 47(2):479-491 (1997); see Attachment 2. For example, Stackenbrandt et al. show that the genus *Micromonosporaceae* have a pattern of 16S rDNA signatures consisting of nucleotides at positions 66-103 (G-C), 127-234 (A-U), 153-168 (C-G), 502-543 (G-C), 589-650 (C-G), etc. (page 487, left hand column, first paragraph). Further to this, Stackenbrandt et al., *supra*, describe phylogenetic studies based on alignments obtained from the RDP (see page 480, left hand column, under Materials and Methods).

Hence, the use of rRNA as molecular chronometers or to determine phylogeny has been used in the art for many years. It is standard in the art to define, identify or characterize new microbial organisms based on signature sequences as determined from a computational alignment analysis using 16S rRNA sequences. It is standard in the art to use the *E.coli* 16S rRNA gene sequence as the reference sequence for identifying the nucleotide positions of the signature sequences (see page 480, right hand column, first paragraph, of Stackenbrandt et al.). Lastly, it is standard in the art to use the Brosius et al. (1978) *E.coli* 16S rRNA gene sequence

to determine phylogeny with respect to microorganisms(see page 480, right hand column, first paragraph).

New claims 25-30 also meet the written description requirement because they claim identical subject matter as that previously presented in claim 1.

Therefore, Applicants submit that based on the discussion above, Dr. Giovannoni's Declaration under 37 CFR § 1.132, and the Woese and Stackenbrandt et al. references (Attachments 1 and 2), Applicants provided a complete disclosure of sufficiently detailed, relevant identifying characteristics to show the Applicants were in possession of the claimed invention.

Accordingly, withdrawal of the rejection of claims 1, 13 and 17-18 and 20 under 35 U.S.C. §112, first paragraph is respectfully requested.

**B. 35 U.S.C. §112, second paragraph (indefiniteness)**

Claims 1, 13-18 and 20 stand rejected under 35 U.S.C. §112, second paragraph as allegedly being indefinite for failing to particularly point out and distinctly claim the subject matter which Applicant regards as the invention. Applicants respectfully traverse this rejection.

Claims 1 and 18 have been amended and claims 25-30 have been added. Claims 13-16 have been canceled. Claims 1 and 18 have been amended per the suggestion of the Office Action on page 4 (i.e. for claim 1 to include the limitations of claims 13-16, and for claim 18 to include the limitation of "sodium containing" growth medium). As described above, the amendments to the claims and paragraph [0026] and Table 3 of the specification are made to improve their form and clarity without adding new subject matter. The amendments to the

In re Application of:

Fenical et al.

Application No.: 09/991,518

Filed: November 16, 2001

Page 14

PATENT

Attorney Docket No.: UCSD1630-1

claims and specification are supported in the specification as filed and in the previous pending claims. Therefore, the metes and bounds of the claimed invention are clearly set forth.

Accordingly, withdrawal of the rejection of claims 1, 13-18 and 20 under 35 U.S.C. § 112, second paragraph is respectfully requested.

In re Application of:  
Fenical et al.  
Application No.: 09/991,518  
Filed: November 16, 2001  
Page 15

PATENT  
Attorney Docket No.: UCSD1630-1


#### IV. Conclusion

Applicants submit that the pending claims are in condition for allowance. Reexamination, reconsideration, withdrawal of the rejections, and early indication of allowance are requested respectfully. If any questions remain, the Examiner is urged to contact the undersigned below.

No fee is believed due in connection with this Amendment. If any additional fees are due, the Commissioner is hereby authorized to charge any fees that may be required by this paper to Deposit Account No. 07-1896. A duplicate copy of this Transmittal Sheet is attached.

Respectfully submitted,

Date: May 20, 2005

  
\_\_\_\_\_  
Lisa A. Haile, J.D., Ph.D.  
Registration No. 38,347  
Telephone: (858) 677-1456  
Facsimile: (858) 677-1465

DLA PIPER RUDNICK GRAY CARY US LLP  
4365 Executive Drive, Suite 1100  
San Diego, California 92121-2133  
**USPTO CUSTOMER NO. 28213**

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

Applicants:	Fenical et al.	Art Unit:	1651
Application No.:	09/991,518	Examiner:	I. Marx
Filing Date:	November 16, 2001	Confirmation No.:	7755
Title:	MARINE ACTINOMYCETE TAXON FOR DRUG AND FERMENTATION PRODUCT DISCOVERY		

Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313

**DECLARATION OF**  
**APPLICANT UNDER 37 C.F.R. § 1.132**

Sir:

I, Stephen J. Giovannoni, do hereby declare and state that:

1. I am a Professor at Oregon State University in Corvallis, Oregon. I conduct research and have published many articles relating to microbial identification and classification, including those based on 16S rRNA alignment and analysis. Currently, I am teaching a course titled, "Microbial Diversity," which relates directly to my area of expertise and interest, as well as to the subject matter of the above-identified patent application. My *curriculum vitae* is attached as Exhibit A.

2. I am on the Scientific Advisory Board of Nereus Pharmaceuticals, located at 10480 Wateridge Circle, San Diego, California 92121. Nereus Pharmaceuticals is the exclusive licensee of the above-identified matter.

3. I understand that claims 1, 13, 17-18 and 20 have been rejected under 35 U.S.C. § 112, first paragraph, in the above-identified patent application, for allegedly failing to comply with the written description requirement. I understand that the alleged failure to satisfy the written description requirement means that the claimed subject matter was allegedly not described in the application in such a way as to reasonably convey to one skilled in the relevant

GT6446056.1  
328342-202

Application No.: 09/991,518  
Applicant: Fenical et al.  
Filed: November 16, 2001  
Page 2

PATENT  
Attorney Docket No.: UCSD1630-1

art, such as myself, that the inventor(s), at the time the application was filed, had possession of the claimed invention.

4. First, as "one skilled in the art," I declare that that the universal "gold standard" for microbial classification is a computational alignment analysis against existing 16S rRNA gene sequences. This same "gold standard" was utilized in the above-identified patent application to identify and characterize the signature sequences in the *Salinospora* isolates.

5. Second, Bergey's Manual of Systematic Bacteriology (George M. Garrity, David R. Boone, Editors) is a text used worldwide, and supports the above mentioned "gold standard" (i.e. microbial identification and classification schemes rely on 16S rRNA sequences). Bergey's Manual of Systematic Bacteriology, is also the taxonomic hierarchy used in the Ribosomal Database Project (RDP). The RDP provides aligned and annotated rRNA gene sequences, and is made available through the RDP website (<http://rdp.cmc.msu.edu/>). See also Cole et al, "The ribosomal database project (RDP-II): sequences and tools for high-throughput rRNA analysis," *Nucleic Acids Research*, 33:D294-296 (2005), which is hereby attached as Exhibit B. The signature nucleotides of the *Salinospora* 16S rRNA sequences claimed in the above identified application were aligned against other 16S rRNA gene sequences, including members of *Micromonosporaceae* in the RDP (see paragraph [0028] and Table 3 of the specification).

6. Third, Brosius et al. (1978) was first to publish the complete 16S rRNA of *E.coli* (please see Brosius et al., "Complete nucleotide sequence of 16S ribosomal RNA gene from *Escherichia coli*," *Proc. Natl. Acad. Sci. USA* 75(1):4801-4805, which is hereby attached as Exhibit C). The Brosius 16S rRNA *E.coli* gene sequence is the reference sequence used worldwide and is the sequence found on the RDP. This is true regardless of any changes and updates to the sequence. To put it another way, the *E.coli* 16S rRNA gene sequence on the RDP is "permanent and invariant." Therefore, the Applicants mention of "*E.coli* positions 27-1492" is a reference to the Brosius' *E.coli* 16S rRNA gene sequence. Those skilled in the art will immediately understand that the Applicants were in possession of the claimed subject matter based on a comparison of the invention *Salinospora* 16S rRNA sequences with *E.coli* 16S rRNA sequence.

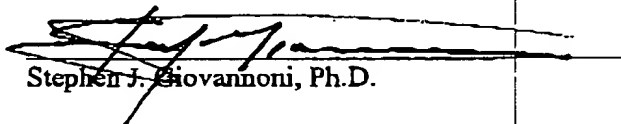
GT6446056.1  
328342-202

Application No.: 09/991,518  
Applicant: Fenical et al.  
Filed: November 16, 2001  
Page 3

PATENT  
Attorney Docket No.: UCSD1630-1

7. I further declare that all statements made herein of knowledge are true and that all statements made on information and belief are believed to be true, and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine, or imprisonment, or both under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

5/17/05  
Date

  
Stephen J. Giovannoni, Ph.D.

GT6446056.1  
328342-202



# **CURRICULUM VITAE**

Stephen J. Giovannoni

Department of Microbiology  
Oregon State University  
Corvallis, OR 97331-3804 U.S.A.

Telephone: (541) 737-1835  
Fax: (541) 737-0496  
E-mail: [steve.giovannoni@oregonstate.edu](mailto:steve.giovannoni@oregonstate.edu)

## **Education:**

- University of Oregon; Ph.D. in Biology 1984
- Boston University; M.A. in Biology 1978
- University of California, San Diego; B.A. in Biology 1974

## **Research Interests:**

- Oceanic and Freshwater Bacterioplankton
- Microbial Genomics
- Microbial Life in the Oceanic Lithosphere

## **Professional Experience:**

- Director, Molecular and Cellular Biology Program,  
Oregon State University, Corvallis. 2000-present
- Professor, Department of Microbiology,  
Oregon State University, Corvallis. 1999-present
- Associate Professor, Department of Microbiology,  
Oregon State University, Corvallis. 1993-1999
- Assistant Professor, Department of Microbiology,  
Oregon State University, Corvallis. 1988-1993
- NSF Postdoctoral Research Fellow with Norman Pace,  
Indiana University, Bloomington. 1984-1988
- Instructor, Department of Biology, University of Oregon, Eugene. 1984
- Graduate Research Assistant with Richard Castenholz,  
Department of Biology, University of Oregon, Eugene. 1979-1984
- Research Associate with Edward Leadbetter, Biological Sciences  
Department, University of Connecticut. 1978-1979
- Graduate Teaching Fellow with Lynn Margulis, Department of  
Biology, Boston University, Boston. 1975-1978
- Research Associate with George Feher, Department of Physics,  
University of California, San Diego. 1974-1975

## **Honors and Awards:**

- Pernot Endowed Professor, OSU Department of Microbiology 2005
- Milton Harris Award for Exceptional Achievement in Microbiology  
College of Science, Oregon State University 2003
- Fellow, American Academy of Microbiology 1997
- Sugihara Young Faculty Research Award,  
College of Science, Oregon State University 1994
- Emerging Scholar Award, Phi Kappa Phi 1993
- Morgenroth Award for Exceptional Achievement as a Graduate Student  
University of Oregon 1984

## **Recent Teaching Activity at Oregon State University:**

- Full responsibility for *Genomics and Cellular Evolution* (MB668), yearly.
- Full responsibility for *Microbial Diversity* (MB420/520), alternate years.
- Full responsibility for *Microbial Diversity Laboratory* (MB421/521), alternate years 1989
- Fall, '03, MCB511 *Research Perspectives* (1.5 lecture hours)
- Fall, '03, MCB 526 *Advanced Biotechnology Techniques* (3 lecture hours)
- Fall, '03, OC669 *Principles of the Subsurface Biosphere* (2 lecture hours)

#### Graduate Training:

- Major or co-advisor for five completed Ph.D.s and six masters degrees.
- Service on over 30 graduate committees.

#### Teaching (other institutions):

- Co-instructor in *Marine Genomics*, The Bermuda Biological Station for Research, July 2003, 2004.
- Lecturer in *Microbial Diversity*, The Rockefeller University, Jan., 1998 and Jan. 2000.
- Co-instructor in *Marine Microbial Ecology*, The Bermuda Biological Station for Research, July 1999-2002.
- Lecturer in *Microbial Phylogeny: Linkages to Processes and Biogeochemistry* (Microbiology 670/470), University of Tennessee, Feb. 1998.
- Instructor, Module Organizer, University of Southern California/ONR. Advanced Techniques Course *Molecular Biology and Biochemistry of Marine Organisms*, July 1992.
- Instructor, Marine Biological Laboratory (Woods Hole, MA) course *Molecular Probes in Marine Ecology*, summer, 1989.

#### Public Outreach:

- Advisor for American of Microbiology/Public Broadcasting  
Production: "Intimate Strangers, Unseen Life on Earth" 1997-1998
- Member of Microbial Literacy Collaborative, an American Society  
for Microbiology organization dedicated to disseminating knowledge  
about microbiology to the general public 1997

#### Invited talks (recent):

- Opening Lecture, Tenth International Symposium on Microbial Ecology, Cancun, Mexico, Aug., '04, *Winogradsky Revisited*.
- Invited Lecture, Ecological Society of America Annual Meeting, Portland, OR. Aug. '04, *Temporal and spatial structure in bacterioplankton populations in the western Sargasso Sea*.
- Invited Lecture, Census of Marine Life Workshop, Monterey, CA, Nov. '03. *High Throughput Methods for the Cultivation of Marine Bacteria*.
- Invited Lecture, ASM Annual Meeting, Washington D.C., May, '03, *The Uncultured Microbial Majority*.
- Invited Lecture, ASM Annual Meeting, Washington D.C., May, '03, *A Targeted High-throughput cultivation strategy*.

#### University Service:

- Director, Molecular and Cellular Biology Program, 2001-2004

#### University Committees:

- Chair of search committee for director of the Center for Gene Research and Biotechnology, 1999-2000.
- Center for Gene Research and Biotechnology Advisory Board member and representative to Research Office, 1997 to present.
- Vice Provost for Research Search Committee member, 1996-1997.
- Rice Endowed Chair in Entomology Search and Oversight Committee member, 1997 to present.

- Chair, OSU Research Council, 1994-1995.
- OSU Research Council member, 1993-1994.

#### Sponsored Seminars and Symposia:

- Co-organizer of Center for Gene Research and Biotechnology Annual Retreat, Sept. 20-21, 1996.
- Organized annual banquet (Chair of Banquet Committee), Phi Kappa Phi., 1997.
- Organized Sugihara symposium *Microbial Diversity* at Oregon State University, Feb. 1996.
- Initiated and organized departmental seminar series, including visit and seminar by Fran Paerler of New England Biolabs, 1996-1997.
- Organized Department of Microbiology Summer Research Symposium, 1996.

#### Professional Societies:

- American Society for Microbiology
- American Association for the Advancement of Science
- American Society of Limnology and Oceanography

#### Professional Service:

- |  |              |
|--|--------------|
| • Associate Editor, <i>Environmental Microbiology</i>  | 2000-present |
| • Editorial Board of <i>Applied and Environmental Microbiology</i>   | 1997-1999    |
| • Grant Panel member, NSF Ecology division   | 1998         |
| • Member of Ocean Drilling Program Deep Biosphere Planning Group   | 1997-1998    |
| • Chair for the Division of Systematic and Evolutionary Biology, American Society for Microbiology             | 1990         |
| • Chair Elect for the Division of Systematic and Evolutionary Biology of the American Society for Microbiology | 1989         |

#### Grant Panelist (Recent):

- Grant Review Panel: NSF/USDA Small Genomes, March, '03
- Grant Review Panel: NOAA Ocean Sciences, Dec. '03.

**Ad-hoc Manuscript Reviews:** *Nature*, *Marine Ecology Progress Series*, *Science*, *International Journal of Systematic Bacteriology*, *Limnology and Oceanography*, *Applied and Environmental Microbiology*, *Proceedings of the National Academy of Sciences U.S.A.*

**Ad-hoc Proposal Reviews:** FASEB, NSF, U. S. Environmental Protection Agency, U. S. Dept. of Agriculture, Australian Research Council.

**Patents:** U.S. No. 10/053,243. High Throughput Microbial Culturing. Pending

#### Research Grants:

- |           |   |
|-----------|---|
| 1986-1988 | NSF Postdoctoral Fellowship Award, Division of Biotic Systems and Resources: "Phylogenetic Analysis of Marine Picoplankton by rRNA Gene Cloning and Sequencing"                     |
| 1988      | OSU Research Council Grant: "Chloroplast Phylogeny by 16S Ribosomal RNA Gene Sequence Analysis," \$4,000.   |
| 1988-1989 | Oregon Medical Research Foundation Grant: "Molecular Phylogeny of Two Protozoan Pathogens: <i>Pneumocystis carinii</i> and <i>Leishmania</i> sp.," \$10,953.                        |
| 1989-1990 | National Science Foundation Grant: " <i>In situ</i> Analyses of the Distributions and Phylogeny of Cultivable and Non-cultivable Planctomycetales Using Phylogenetic Group-Specific |

RNA Probes," BSR-8818167, \$110,000.

- 1989-1995 National Dairy Promotion and Research Board Grant: "Probes for Conserved 16S Ribosomal RNA (rRNA) Gene Sequences to Isolate *Lactococcus cremoris* from Nature," \$210,000.
- 1991 National Science Foundation Grant: "*In situ* Analyses of the Distributions and Phylogeny of Cultivable and Non-cultivable *Planctomycetales* Using Phylogenetic Group-Specific rRNA Probes," BSR-9020477, \$75,000.
- 1991 OSU Research Council Grant: "*In Situ* Microscopic Quantification of Low-copy Number Ribosomal RNA Targets by SIT Camera Image Analysis," \$4,000.
- 1991-1993 National Science Foundation Grant: "Molecular Analyses of the Population Dynamics and Activity of a Newly Identified Bacterioplankton Group," OCE-9016373, \$368,460.
- 1992 Oregon Advanced Computing Institute Grant: "gRNAid: An Interactive Graphics Program for Predicting the Secondary Structures of Ribonucleic Acid Molecules," 40-0140, \$19,315.
- 1993-1997 Department of Energy, Ocean Margins Program Grant: "The Dynamics of Carbon Exchange in Vertically Stratified Coastal Bacterioplankton Communities." FG0693ER61697, \$386,189.
- 1994 Joint Oceanographic Institutions Grant: "Genetic Evidence for Endolithic Microbial Life colonizing Basaltic glass/seawater interfaces"; Co-P.I. with M. Fisk. \$12,000.
- 1995-1998 National Science Foundation Grant: "Antarctic Lake Ice Microbial Consortia: Origin, Distribution and Growth Physiology", PP-9419423, \$108,327 (to S.J.G.).
- 1995-1996 OSU Research Council Grant: The Kinetics of Gene Amplification and Chimera Formation in the Polymerase Chain Reaction", \$7,973.
- 1997-1999 National Science Foundation Grant: "Evidence for Endolithic Microbes in Oceanic Basalts", Co-P.I. with M. Fisk, OCE-9618728, \$38,350 (to S.J.G.).
- 1997-2000 National Science Foundation Grant: "Interactions Between Bacterioplankton Communities and Dissolved Substrates at the Bermuda Atlantic Time Series Study Station", OCE-9618530, with C. Carlson, co.-P.I., \$293,000 (to S.J.G.).
- 1997-2000 National Science Foundation Grant: "Spatial, Temporal and Phylogenetic Structure of Bacterioplankton Communities in Crater Lake, Oregon", DEB-9709012, with E. Urbach, co-P.I., \$303,588.
- 1998-1999 National Science Foundation Grant: "Development of Capability to Measure Proxies of Microbial Activity Within Ocean Crust", BES-9729672, with James Cowen (P.I.), F. Kenig and H.P. Johnson. \$54,881 (to S.J.G.).
- 1998-2001 National Science Foundation Grant: "Time-series Responses to a Mid-Ocean Ridge Volcanic Event: Juan de Fuca and Gorda Ridges", OCE-9902048, with James Cowen (P.I.). \$94,298 (to S.J.G.).
- 1998-1999 National Science Foundation Grant: Quantitative Imaging of the Smallest Bacterioplankton Cells, SAR11, at the Theoretical Limits of Light Microscopy. OCE-9816489. \$19,500.

- 1999-2001 Collaborative Research on Bacterioplankton Biology and Biochemistry at the Bermuda Atlantic Time-series Station: An Oceanic Microbial Observatory. MCB-9977930. \$299,990 (to S.J.G.).
- 1999-2001 National Science Foundation Major Research Instrumentation Grant: Advanced Microbe Isolation Laboratory. OIA-9977469. \$338,940.
- 1999-2001 Murdock Charitable Trust Grant: Microbe Discovery by Solid State Cytometry with Fluorescent DNA Probes. \$306,730.
- 2000-2002 Oregon Sea Grant: Are Algicidal Bacterial Important in Controlling Phytoplankton Blooms in Oregon Coastal Waters? R/HAB-01. \$236,564.
- 2000-2003 National Science Foundation Grant: Effects of Microbial Activity on Rates of Basalt Alteration. With co-P.I. M Fisk. OCE-0085436. \$358,999.
- 2001 Diversa Corporation Contract: High Throughput Culturing. \$225,000.
- 2001-2006 National Science Foundation Proposal. IGERT- The Earth's Subsurface Biosphere. Co-PI. With M. Fisk. \$2,674,860.
- 2001-2004 Sloan Foundation Grant: Professional Masters Degree Programs in the Sciences at Oregon State University. \$400,924.
- 2002-2005 National Science Foundation Grant: Coastal Bacterioplankton Systematics: A High Throughput Culturing Approach. DEB 0207085. \$275,999.
- 2003-2008 National Science Foundation Grant: Collaborative Research Linking Microbial Discovery to Biogeochemical Processes: An Oligotrophic Oceanic Microbial Observatory. MCB-0237713. \$306,102.
- 2003-2004 National Science Foundation Grant: Collaborative Research: Microbial Diversity and Function in the Permanently Ice-covered Lakes of the McMurdo Dry Valleys, Antarctica. \$99,989.
- 2005-2010 Gordon and Betty Moore Foundation: Genomics of Oceanic Bacteria. \$3,640,000.

**Peer Reviewed Articles (published, in press, submitted, or near submission, in reverse order):**

83. Giovannoni, S.J., L. Bibbs, S. Givan, J. Tripp, M. Podar, M. Noordewier, J. Eads, E. Mathur, M. S. Rappé, and K. L. Vergin. Genome Streamlining in a Cosmopolitan Oceanic Bacterium. In Preparation.
82. Cho, J. C., M. D. Staples, R. M. Morris, K. L. Vergin, D. F. Barofsky, and S. J. Giovannoni. Proteomics links novel aerobic anoxygenic phototroph isolates to the oceanic metagenome. In review.
81. Giovannoni, S.J., D. Barofsky, L. Bibbs, J.C. Cho, R. Desiderio, S. Laney, E. Mathur, M. S. Rappé, M. Staples, and K. L. Vergin. Proteorhodopsin phototrophy in the Ubiquitous Marine Bacterium *Pelagibacter ubique*. In review.
80. Morris, R.M., J.C. Cho, M.S. Rappé, K. L. Vergin, C. A. Carlson, and S. J. Giovannoni. Bacterioplankton responses to deep season mixing in the Sargasso Sea. *Limnol. Oceanog.* In review.

79. Page, K.A., S. A. Connon, and S.J. Giovannoni. 2004. Oligotrophic isolates from Crater Lake, Oregon are representative of dominant freshwater bacterioplankton. 70:6542-6550.
78. Staples, M.D. J. C. Cho, S. J. Giovannoni and D.F. Barofsky. 2004. Proteomic analysis of novel marine bacteria using MALDI and ESI spectrometry. J. Biomolec. Techniques. 15:191-198.
77. Connon, S.A., A. Tovanabootr, M. Dolan, K. Vergin, S.J Giovannoni, and L. Semprini. 2004. Bacterial Community Composition Determined by Culture Independent and Dependant Methods during Propane Stimulated Bioremediation in Trichloroethene Contaminated Groundwater. Environ. Microbiol. 7:165-78.
76. Carlson, C.A., S. J. Giovannoni, D. A. Hansell, S. J. Goldberg, R. Parsons, K. Vergin. 2004. Interactions between DOC, microbial processes, and community structure in the mesopelagic zone of the northwestern Sargasso Sea. Limnol. Oceanog. 49:1073-1083.
75. Morris, R. M., M. S. Rappé, E. Urbach, S. A. Connon, S. J. Giovannoni. 2004. Prevalence of the Chloroflexi-Related SAR202 Bacterioplankton Cluster Throughout the Mesopelagic Zone and Deep Ocean. Environ. Microbiol. 70:2836-2842.
74. Cho, J. C., and S. J. Giovannoni. 2004. *Robiginitalea biformata* gen. nov., sp. nov., a new marine bacterium in the family Flavobacteriaceae that contains higher G+C composition. Int. J. Syst. Evol. Microbiol. 54:1101-1106.
73. Cho, J. C., and S. J. Giovannoni. 2004. *Oceanicola granulosus* gen. nov., sp. nov. and *Oceanicola batsensis* sp. nov., poly-beta-hydroxybutyrate-producing marine bacteria in the order "Rhodobacterales". Int. J. Syst. Evol. Microbiol. 54:1129-1136.
72. Cho, J. C., K. L. Vergin, R. M. Morris, and S. J. Giovannoni. 2004. Discovery of the novel bacterial phylum Lentisphaerae with cultivation of *Lentisphaera araneosa* gen. nov., sp. nov, a transparent exopolymer producing marine bacterium. Environ. Microbiol. 6:611-621.
71. Cho, J. C., and S. J. Giovannoni. 2004. Cultivation and growth characteristics of a diverse group of oligotrophic marine Gammaproteobacteria. Appl. Environ. Microbiol. 70:432-440.
70. Cho, J.C., and S. J. Giovannoni. 2003. *Fulvimarina pelagi* gen. nov., sp. nov., a marine bacterium that forms a deep evolutionary lineage of descent in the order "Rhizobiales". Int. J. Syst. Evol. Microbiol. 53:1853-1859.
69. Cho, J.-C., and S. J. Giovannoni. 2003. *Parvularcula bermudensis* gen. nov., sp. nov., a marine bacterium that forms a deep branch in the  $\alpha$ -Proteobacteria. Int. J. Syst. Evol. Microbiol. 53:1031-1036.
68. Cho, J.-C., and S. J. Giovannoni. 2003. *Croceibacter atlanticus* gen. nov., sp. nov., a novel marine bacterium in the family Flavobacteriaceae. Syst. Appl. Microbiol. 26:76-83.
67. Cowen J., S.J. Giovannoni, H. P. Johnson, F. Kenig, D. Butterfield, M. Rappé, Hutnak, and P. Lam. 2003. Fluids from aging ocean crust that support microbial life. Science 299:120-123.
66. Morris, R.M., M. S. Rappé, S. A. Connon, K. L. Vergin, W. A. Siebold, C.A. Carlson, S. J. Giovannoni. 2002. High Cellular Abundance of the SAR11 Bacterioplankton Clade in Seawater. Nature 420: 806-810.
65. Rappé, M.S , S. Connon, S., K. L. Vergin, and S. J. Giovannoni. 2002. Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. Nature 418:630-631.

64. Carlson, C.A., S. J. Giovannoni, D. A. Hansell, S. J. Goldberg, R. Parsons, M. P. Otero, K. Vergin, and B. R. Wheeler. 2002. The effect of nutrient amendments on bacterioplankton production, community structure, and DOC utilization in the northwestern Sargasso Sea. *Aquatic Microbial Ecology*. 30:19-36.
63. Connon, S and S.J Giovannoni. 2002. High throughput methods for culturing microorganisms in very low nutrient media yield diverse new marine isolates. *Appl. Environ. Microbiol.* 68: 3878-3885.
62. Vergin, K. L., M.S. Rappé, , and S.J. Giovannoni. 2001. Streamlined method to analyze 16S rRNA gene clone libraries. *Biotechniques* 30:938-944.
61. Urbach, E., K. L. Vergin, L. Young, A. Morse, G. Larson and S.J. Giovannoni. 2001. Unusual bacterioplankton in Crater Lake Oregon. *Limnol. Oceanog.* 46:557-572.
60. Lanoil, B. D., C. Carlson and S.J. Giovannoni. 2000. Bacterial chromosomal painting for in situ monitoring of cultured marine bacteria. *Environ. Microbiol.* 2:654-665.
59. Rappé, M.S., Vergin, K. and Giovannoni, S.J. 2000 Phylogenetic comparisons of a coastal bacterioplankton community with its counterparts in open ocean and freshwater systems. *FEMS Microb Ecol* 33: 219-232.
58. Gordon, D.A., J. Priscu and S.J. Giovannoni, 2000. Origin and phylogeny of microbes living in permanent Antarctic lake ice. *Microb. Ecol.* 39:197-202.
57. Janson, S., Bergman, B., Carpenter, E.J., Giovannoni, S.J., Vergin, K. 1999. Genetic Analysis of the Marine Diazotrophic cyanobacterium *Trichodesmium*. *FEMS Microb. Ecol.* 30:57-65.
56. Fisk, M. and S.J. Giovannoni. Sufficient conditions for a deep biosphere on Mars. 1999. *Journal of Geophysical Research, Planets.* 104:11,805-11,815.
55. Urbach, E., K. L. Vergin and S.J. Giovannoni. 1999. Immunochemical detection and isolation of DNA from metabolically active bacteria. *Appl. Environ. Microbiol.* 65:1207-1213.
54. McAshan, S.K., K.L. Vergin, S.J. Giovannoni and D.S. Thaler. 1999. Interspecies hybridization in the Enterococci via conjugation of chromosomal vancomycin resistance. *Microb. Drug Resis.* 5:101-112.
53. Mauel, M.J., S.J. Giovannoni and J.L. Fryer. 1999. Phylogenetic analysis of *Piscirickettsia salmonis* isolates by 16S ribosomal DNA sequencing. *Dis. Aquat. Org.* 35:115-123.
52. Rappé, M.S., D.A. Gordon, K. L. Vergin, S.J. Giovannoni. 1999. Phylogeny of Actinobacteria-related SSU rRNA gene clones recovered from marine bacterioplankton. *Syst. Appl. Microbiol.* 22:106-112.
51. Suzuki, M., M.S Rappé, and S.J. Giovannoni. 1998. Kinetic bias in estimates of coastal picoplankton community structure obtained by measurements of SSU rDNA PCR-amplicon length heterogeneity. *Appl. Environ. Microbiol.* 64:4522-4529.
50. Fisk, M. R., S.J. Giovannoni and I. Thorseth. 1998. Alteration of oceanic volcanic glass: textural evidence for microbial activity. *Science* 281: 978-980.
49. Vergin, K., E. Urbach, J. L. Stein, E. F. DeLong and S.J. Giovannoni. 1998. Screening of a fosmid library of marine environmental genomic DNA fragments reveals four clones related to Planctomycetales. *Appl. Environ. Microbiol.* 64: 3075-3078.

48. Urbach, E., C. Schindler and **S.J. Giovannoni**. 1998. A PCR fingerprinting technique to distinguish isolates of *Lactococcus lactis*. FEMS Microbiol. Let. 162:111-115.
47. Rappé, M.S., M. Suzuki, K. L. Vergin, **S.J. Giovannoni**. 1998. Phylogenetic diversity of ultraplankton plastid SSU rRNA genes recovered in environmental nucleic acid samples from the Pacific and Atlantic coasts of the United States. Appl. Environ. Microbiol. 64:294-303.
46. Priscu, J., C.H. Fritsen, E. Adams, **S.J. Giovannoni**, H. Paerl, C. McKay, D. Gordon, and B. Lanoil. 1998. Perennial Antarctic lake ice: an oasis for life in a polar desert. Science 280:2095-2098.
45. Wright, T. D., K. Vergin, P. Boyd and **S.J. Giovannoni**. 1997. A novel  $\beta$ -Proteobacterial lineage from the lower ocean surface layer. Appl. Environ. Microbiol. 63:983-989.
44. Lanoil, B. D., and **S.J. Giovannoni**. 1997. Identification of bacterial cells by chromosomal painting. Appl. Environ. Microbiol. 63:1118-1123.
43. Suzuki, M., M. S. Rappé, Z.W. Haimberger, H. Winfield, N. Adair, J. Ströbel and **S.J. Giovannoni**. 1997. Bacterial diversity among SSU rDNA gene clones and cellular clones from the same seawater sample. Appl. Environ. Microbiol. 63:983-989.
42. Urbach, E., B. Daniels, M. S. Salama W. E. Sandine and **S.J. Giovannoni**. 1997. The *ldh* phylogeny for environmental isolates of *Lactococcus lactis* is consistent with the rRNA genotypes, but not with phenotypes. Appl. Environ. Microbiol. 63: 694-702.
41. Field, K.G., N. Adair, D.A. Gordon, M. S. Rappé and **S.J. Giovannoni**. 1997. Genetic diversity and depth-specific speciation within the SAR11 cluster, a marine bacterial lineage. Appl. Environ. Microbiol. 63:63-70.
40. Rappé, M.S., P.F. Kemp and **S.J. Giovannoni**. 1997. Phylogenetic diversity of marine coastal picoplankton 16S rRNA genes cloned from the continental shelf off Cape Hatteras, N.C. Limnol. Oceanog. 42:811-826.
39. Mauel, M.J., **S.J. Giovannoni** and J.L. Fryer. 1997. Development of polymerase chain reaction assays for detection, identification and differentiation of *Piscirickettsia salmonis*. Dis. Aquat. Org. 26:189-195.
38. Lanoil, B. D., L.M. Ciuffettii and **S.J. Giovannoni**. 1996. The marine bacterium *Pseudalteromonas haloplanktis* has a complex genome structure composed of two separate genetic units. Genome Res. 6:1160-1169.
37. **Giovannoni, S.J.**, M. S. Rappé, K. L. Vergin and N. Adair. 1996. 16S rRNA genes reveal stratified open ocean bacterioplankton populations related to the Green Non-Sulfur bacteria. Proc. Natl. Acad. Sci. U.S.A. 93:7979-7984.
36. Gordon, D.A. and **S.J. Giovannoni**. 1996. Stratified microbial populations related to *Chlorobium* and *Fibrobacter* detected in the Atlantic and Pacific oceans. Appl. Environ. Microbiol. 62:1171-1177.
35. Suzuki, M., and **S. J. Giovannoni**. 1996. Bias caused by template annealing in the amplification of 16S rRNA genes by PCR. Appl. Environ. Microbiol. 62:625-630.
34. **Giovannoni, S.J.**, M. R. Fisk, Mullins, T.D. and Furnes, H. 1996. Genetic evidence for endolithic microbial life colonizing basaltic glass/seawater interfaces. Proceedings of the Ocean Drilling Program 148:207-214.



33. Rappe', M. S., Kemp, P. F., and S.J. Giovannoni. 1995. Chromophyte plastid 16S ribosomal RNA genes found in a clone library from Atlantic Ocean seawater. *J. Phycol.* 31:979-988
32. Salama, M. S., Musafija-Jeknic, T., W. E. Sandine, and S. J. Giovannoni. 1995. An ecological study of lactic acid bacteria: isolation of new strains of *Lactococcus* including *Lactococcus lactis* subspecies *cremoris*. *Journal of Dairy Science* 78:1-14.
31. Salama, S., W. Sandine, and S. J. Giovannoni. 1995. A milk-based method for detecting antimicrobial substances produced by lactic acid bacteria. *J. Dairy Sci.* 78:1219-1223.
30. Mullins, T. D., T. B. Britschgi, R. L. Krest, and S. J. Giovannoni. 1995. Genetic comparisons reveal the same unknown lineages in Atlantic and Pacific bacterioplankton communities. *Limnol. Oceanog.* 40:148-158.
29. Salama, S., W. Sandine, and S. J. Giovannoni. 1993. Isolation of *Lactococcus lactis* subsp. *cremoris*. from nature by colony hybridization with rRNA probes. *Appl. Environ. Microbiol.* 57:1313-1318.
28. Lovley, D. R., S. J. Giovannoni, D. C. White, J. E. Champine, E. Phillips, Y. A. Gorby, and S. Goodwin. 1993. *Geobacter metallireducens* gen. nov. sp. nov., a microorganism capable of coupling the complete oxidation of organic compounds to the reduction of iron and other metals. *Archive Microbiol.* 159:336-344.
27. Cary, S. C. and S. J. Giovannoni. 1993. Transovarial inheritance of endosymbiotic bacteria in deep-sea vesicomyid clams. *Proc. Natl. Acad. Sci. USA* 90:5695-5699.
26. Cary, S. C., W. Warren, E. Anderson, and S. J. Giovannoni. 1993. Identification and localization of bacterial endosymbionts in hydrothermal vent taxa with symbiont-specific PCR amplification and *in situ* hybridization techniques. *Mol. Mar. Biol. Biotech.* 2:251-262.
25. Liesack, W., R. Soller, T. Stewart, H. Haas, S. J. Giovannoni, and E. Stackebrandt. 1992. The influence of tachytelically (rapidly) evolving sequences on the topology of phylogenetic trees -intrafamily relationships and the phylogenetic position of the Planctomycetaceae as revealed by comparative analysis of 16S ribosomal RNA sequences. *System. Appl. Microbiol.* 15:357-362.
24. Lane, D. J., A. P. Harrison, Jr., D. Stahl, B. Pace, S. J. Giovannoni, G. J. Olsen, and N. R. Pace. 1992. Evolutionary relationships among sulfur- and iron-oxidizing eubacteria. *J. Bacteriol.* 174(1):269-278.
23. Fryer, J. L., C. N. Lannan, S. J. Giovannoni, and N. D. Wood. 1992. *Pisericettsia salmonis* gen. nov., sp. nov., the causative agent of an epizootic disease in salmonid fishes. *Int. J. Sys. Bacteriol.* 42:120-126.
22. Field, K. G., S. M. Landfear, and S. J. Giovannoni. 1991. 18S rRNA sequences of *Leishmania enrietti* promastigote and amastigote. *International Journal for Parasitology* 21:483-485.
21. Britschgi, T. B. and S. J. Giovannoni. 1991. Phylogenetic analysis of a natural marine bacterioplankton population by rRNA gene cloning and sequencing. *Appl. Environ. Microbiol.* 57:1707-1713.
20. Salama, S., W. Sandine, and S. J. Giovannoni. 1991. Development and application of oligonucleotide probes for identification of *Lactococcus lactis* subsp. *cremoris*. *Appl. Environ. Microbiol.* 57:1313-1318.
19. Gutenberger, S. K., S. J. Giovannoni, K. G. Field, J. L. Fryer, and J. S. Rohovec. 1991. A phylogenetic comparison of the 16S and rRNA sequence of the fish pathogen, *Renibacterium salmoninarum*, to Gram-

positive bacteria. FEMS Microbiol. Let. 77:151-156.

18. Giovannoni, S. J., E. F. DeLong, T. M. Schmidt, and N. R. Pace. 1990. Tangential flow filtration and preliminary phylogenetic analysis of marine picoplankton. Appl. Environ. Microbiol. 56:2572-2575.

17. Giovannoni, S. J., T. B. Britschgi, C. L. Moyer, and K. G. Field. 1990. Genetic diversity in Sargasso Sea bacterioplankton. Nature 345:60-63.

16. Huss, A. R. and S. J. Giovannoni. 1989. Primary structure of the chloroplast small subunit ribosomal RNA gene from *Chlorella vulgaris*. Nucleic Acids Res. 22:9487.

15. Turner, S., T. Burger-Wiersma, S. J. Giovannoni, L. R. Mur, and N. R. Pace. 1989. The relationship of a prochlorophyte, *Prochlorothrix hollandica*, to green chloroplasts. Nature 337:380-382.

14. Weisburg, W. G., S. J. Giovannoni, and C. R. Woese. 1989. The *Deinococcus* - *Thermus* phylum and the effect of rRNA composition on phylogenetic tree construction. System. Appl. Microbiol. 11:128-134.

13. Bomar, D., S. J. Giovannoni, and E. Stackbrandt. 1988. A unique type of eubacterial 5S rRNA in members of the order Planctomycetales. J. Mol. Evol. 27:121-125.

12. Distel, D. L., D. L. Lane, G. J. Olsen, S. J. Giovannoni, B. Pace, N. Pace, D. Stahl, and H. Felbeck. 1988. Sulfur-oxidizing bacterial endosymbionts: Analysis of phylogeny and specificity by 16S ribosomal RNA sequences. J. Bacteriol. 170:2506-2510.

11. Field, K. G., G. J. Olsen, D. J. Lane, S. J. Giovannoni, M. T. Ghiselin, E. C. Raff, N. R. Pace, and R. A. Raff. 1988. Molecular phylogeny of the animal kingdom based on 18S ribosomal RNA sequences. Science 239:748-753.

10. Giovannoni, S. J., E. DeLong, G. J. Olsen, and N. R. Pace. 1988. Phylogenetic group-specific oligodeoxynucleotide probes for *in situ* microbial identification. J. Bacteriol. 170:720-726.

9. Giovannoni, S. J., S. Turner, G. T. Olsen, S. Barns, D. T. Lane, and N. R. Pace. 1988. Evolutionary relationships among cyanobacteria and green chloroplasts. J. Bacteriol. 170:3584-3592.

8. Karl, D. M., G. T. Taylor, J. A. Novitski, H. W. Jannasch, C. O. Wirsen, N. R. Pace, D. J. Lane, G. J. Olsen, and S. J. Giovannoni. 1988. A microbiological study of Guymas basin high temperature hydrothermal vents. J. Deep Sea Res. 35:777-791.

7. Giovannoni, S. J., E. Schabtach, and R. W. Castenholz. 1987. *Isosphaera pallida*, gen. and comb. nov., a gliding, budding eubacterium from hot springs. Arch. Microbiol. 147:276-284.

6. Giovannoni, S. J., W. Godchaux, E. Schabtach, and R. W. Castenholz. 1987. Cell wall and lipid composition of *Isosphaera pallida*, a budding eubacterium from hot springs. J. Bacteriol. 169:2702-2707.

5. Giovannoni, S. J., D. M. Ward, N. P. Revsbech, and R. W. Castenholz. 1987. Obligately phototrophic *Chloroflexus*: primary production in anaerobic, hot spring microbial mats. Arch. Microbiol. 147:80-87.

4. Pierson, B. K., S. J. Giovannoni, D. L. Stahl, and R. W. Castenholz. 1985. *Heliothrix oregonensis*, gen. nov., sp. nov., a phototrophic filamentous gliding bacterium containing bacteriochlorophyll *a*. Arch. Microbiol. 142:164-167.

3. Pierson, B. K., S. J. Giovannoni, and R. W. Castenholz. 1984. Physiological ecology of a gliding bacterium containing bacteriochlorophyll *a*. Appl. Environ. Microbiol. 47:576-584.

2. Giovannoni, S. J., and L. Margulis. 1981. A red *Benetkea* from Laguna Figueroa, Baja, California. *Microbios* 30:47-63.

1. Margulis, L., E. S. Barghoorn, D. Ashendorf, S. Banerjee, D. Chase, S. Francis, S. J. Giovannoni, and J. Stolz. 1980. The microbial community in the layered sediments at Laguna Figueroa, Baja, California. *Precam. Res.* 11:93-123.

#### Reviews, Book Chapters and Other Non-Peer Reviewed Publications:

14. Giovannoni, S. J. 2004. Oceans of bacteria. *Nature* 430:515-516.

13. M. Rappe' and Giovannoni, S. J. 2003 The Uncultured Microbial Majority. *Ann. Rev. Microbiol.* 57:369-394.

12. Giovannoni, S. J., and M. Rappe'. 2000. Evolution, Diversity and Molecular Ecology of Marine Prokaryotes. p. 47-84. *In* Kirchman, D. (ed.) *Microbial Ecology of the Oceans*. John Wiley & Sons, Inc., New York.

11. Giovannoni, S. and M. Rappe'. 1999. Microbial Diversity: It's a New World. *The NEB Transcript.* 10:1-4.

10. Giovannoni, S. J., M. Rappe', D. Gordon, E. Urbach, M. Suzuki, and K. G. Field. 1996. Ribosomal RNA and the evolution of bacterial diversity. p. 63-85. *In* Roberts, D. McL., Sharp, P. Alderson, G. and Collins, M. (ed.) "Evolution of Microbial Life". Society for General Microbiology Symposium 54. Cambridge University Press.

9. Giovannoni, S. J., T. Mullins, and K. G. Field. 1995. Microbial diversity in marine systems: rRNA approaches to the study of unculturable microbes. *In*: "Molecular Ecology of Aquatic Microbes," ed. Ian Joint, Springer-Verlag, Berlin-Heidelberg-New York-Tokyo.

8. Giovannoni, S. J., and S. C. Cary. 1993. Probing marine systems with ribosomal RNAs. *Oceanography* 6:95-104.

7. Giovannoni, S. J., N. Wood, and V. A. R. Huss. 1993. Molecular Phylogeny of Oxygenic Phototrophic Cells and Organelles from Small-Subunit Ribosomal RNA Sequences. Pages 159-170. *In*: *Origins of Plastids*, R. A. Lewin (ed.) Chapman and Hall, NY, NY.

6. Staley, J. T., J. L. Fuerst, S. Giovannoni, and H. Schlesner. 1991. The Order *Planctomycetales* and the Genera *Planctomyces*, *Pirellula*, *Gemmata* and *Isosphaera*. Pages 3710-3731. *In*: M. Dworkin et al. (eds.) *The Prokaryotes*. Volume 4, Chapter 203. Springer-Verlag, New York.

5. Giovannoni, S. J. 1991. The polymerase chain reaction. Pages 177-203. *In*: E. Stackebrandt, and M. Goodfellow (eds.) *Modern Microbiological Methods: Nucleic Acids Techniques in Bacterial Systematics*. John Wiley and Sons, New York.

4. Giovannoni, S. J. and R. W. Castenholz. 1989. Genus *Isosphaera* Giovannoni and Castenholz. Pages 1959-1961. *In*: Krieg, N. R. and J. G. Holt (eds.) *Bergey's Manual of Systematic Bacteriology*, Volume 2. Williams and Wilkins, Baltimore/London.

3. Raff, R. A., K. G. Field, G. J. Olsen, S. J. Giovannoni, M. T. Ghiselin, D. J. Lane, N. R. Pace, and E. C. Raff. 1989. The phylogeny of the animal kingdom: a molecular approach. *In*: B. Fernholm (ed.) *The*

Hierarchy of Life. 70th Nobel Symposium, Sweden. Elsevier.

2. Turner, S., E. F. DeLong, **S. J. Giovannoni**, G. J. Olsen, and N. R. Pace. 1989. Phylogenetic analysis of microorganisms and natural populations. Pages 390-401. In: Cohen, Y. and E. Rosenberg (eds.) Microbial Mats: Ecological Physiology of Benthic Microbial Communities.

1. Olsen, G. J., D. L. Lane, **S. J. Giovannoni**, N. R. Pace, and D. A. Stahl. 1986. Microbial ecology and evolution: a ribosomal RNA approach. Ann. Rev. Microbiol. 40:337-366.

# The Ribosomal Database Project (RDP-II): sequences and tools for high-throughput rRNA analysis

J. R. Cole<sup>1,\*</sup>, B. Chai<sup>1</sup>, R. J. Farris<sup>1</sup>, Q. Wang<sup>1</sup>, S. A. Kulam<sup>1</sup>, D. M. McGarrell<sup>1</sup>,  
G. M. Garrity<sup>1,2</sup> and J. M. Tiedje<sup>1,2</sup>

<sup>1</sup>Center for Microbial Ecology and <sup>2</sup>Department of Microbiology and Molecular Genetics,  
Biomedical Physical Sciences, Michigan State University, East Lansing, MI 48824-4320, USA

Received September 14, 2004; Revised and Accepted September 24, 2004

## ABSTRACT

The Ribosomal Database Project (RDP-II) provides the research community with aligned and annotated rRNA gene sequences, along with analysis services and a phylogenetically consistent taxonomic framework for these data. Updated monthly, these services are made available through the RDP-II website (<http://rdp.cme.msu.edu/>). RDP-II release 9.21 (August 2004) contains 101 632 bacterial small subunit rRNA gene sequences in aligned and annotated format. High-throughput tools for initial taxonomic placement, identification of related sequences, probe and primer testing, data navigation and subalignment download are provided. The RDP-II email address for questions or comments is [rdpstaff@msu.edu](mailto:rdpstaff@msu.edu).

## DESCRIPTION

Release 9 introduces substantial changes to the Ribosomal Database Project (RDP). These changes are in response to the rapidly increasing number of available ribosomal RNA gene sequences (rRNA sequences) and the trend toward high-throughput rRNA sequencing with the concomitant need for high volume rRNA analysis tools. This paper describes changes since the 2003 description (1). Details about the data and analysis services can be found at the RDP-II website (<http://rdp.cme.msu.edu/>).

**Sequences.** The RDP obtains bacterial rRNA sequences from the International Nucleotide Sequence Databases (INSD: GenBank/EMBL/DDBJ) on a monthly basis. These sequences are aligned against a general bacterial rRNA model using a modified version of RNACAD (2), a Stochastic Context Free Grammar (SCFG)-based rRNA aligner that directly incorporates rRNA secondary structure information

into its internal model. This aligner is trained on a set of high-quality hand-aligned sequences and incorporates the conserved bacterial secondary structure model of Gutell and co-workers (3). As of release 9.21 (August 2004), the database contained 101 632 total small subunit bacterial rRNA sequences. Of these, 39 772 were near full-length ( $\geq 1200$  bases), 54 316 came from uncultured organisms and 4431 were from type strains of validly named bacterial species.

**Taxonomy.** All Release 9 tools use a new hierarchical framework (RDP Hierarchy) differing significantly from the hierarchy provided with previous RDP releases. The RDP Hierarchy is based on the new phylogenetically consistent higher-order bacterial taxonomy proposed by Garrity *et al.* (4) (<http://dx.doi.org/10.1007/bergeysoutline>). This hierarchy provides order to the collection. It provides a phylogenetic framework into which to place results of the RDP analysis functions, and it provides an entry point for users looking for sequences from specific groups of organisms. New sequences are placed into the RDP Hierarchy using the RDP Classifier (see below).

## Analysis services

The RDP analysis services have been completely revised to support the emerging trend toward high-throughput rRNA sequence analysis in microbial ecology and related disciplines. Three of the tools listed below incorporate the concept of data filters. The user can choose to apply up to three data filters on the view or analysis. By applying the three filters, the user can (i) include only environmental clone or only isolate sequences; (ii) include only sequences  $\geq 1200$  bases in length (near full-length) or only shorter sequences; and (iii) include only sequences from type strains or only non-type strain sequences. The latter filter is of special importance since type strains act as a link between rRNA-based phylogeny and taxonomy. A more detailed description of each analysis service can be found at the RDP website.

\*To whom correspondence should be addressed. Tel: +1 517 432 4998; Fax: +1 517 353 8957; Email: [rdpstaff@msu.edu](mailto:rdpstaff@msu.edu)

The online version of this article has been published under an open access model. Users are entitled to use, reproduce, disseminate, or display the open access version of this article for non-commercial purposes provided that: the original authorship is properly and fully attributed; the Journal and Oxford University Press are attributed as the original place of publication with the correct citation details given; if an article is subsequently reproduced or disseminated not in its entirety but only in part or as a derivative work this must be clearly indicated. For commercial re-use permissions, please contact [journals.permissions@oupjournals.org](mailto:journals.permissions@oupjournals.org).

**Hierarchy Browser** allows rapid navigation through the RDP sequence data. The browser presents views of the RDP sequences placed either in the RDP Hierarchy, or optionally in the NCBI taxonomy hierarchy (5). While navigating, the browser automatically expands an appropriate number of hierarchical levels to fit the display. At any time, the user can select for later download of both individual sequences and those of entire taxa. Data filters can be applied at any time to limit the display to specific data subsets. In addition, the user can quickly search for words or phrases in the sequence definition line. This includes the organism name and strain designation (if available), culture collection identifiers and INSD nucleic acid accession identifiers.

**RDP Classifier** places sequences into the RDP Hierarchy. Optimized for large query sets, it can be used to give an initial taxonomic placement for a single sequence or hundreds of sequences. The first result page summarizes the assignments on an interactive display similar to that of the Hierarchy Browser. Each node in the hierarchy lists the number of user queries assigned to that taxonomic rank. A confidence estimate is generated for each assignment, and the assignments are displayed only when the estimate is above a user-specified confidence threshold. At any time, the user can switch to a detail view showing the detailed taxonomic assignments and confidence scores for any subset of query sequences. These assignment details can also be downloaded in a file suitable for import into popular spreadsheet programs.

**Sequence Match** is a complete re-implementation of the original Sequence Match method (1). Sequence Match finds sequences similar to a user's query sequences using a word matching strategy not requiring prior alignment. Sequence Match is more accurate than BLAST (6) at finding closely related rRNA sequences (Table 1). The related sequences returned by Sequence Match serve as a good starting point for more detailed examination of relatedness by classical phylogenetic or other methods. The initial result page presents a *k*-nearest neighbor (*k*-NN) classifier assignment of the query sequences. A query is assigned to the lowest taxonomic rank that includes the *k* highest scoring database sequences. The value of *k*, as well as the three data filters can be changed at will in this view. The user can switch from the summary *k*-NN view to a detailed results view for any query sequence. In this view, the top *k* database matches to the query are displayed in the RDP Hierarchy. In this mode, any subset of the matches can be selected for transfer to the Hierarchy Browser and later

download. A third view presents sets of results in a format suitable for download.

**Probe Match** is a complete re-implementation of our previous Probe Match program (1). It uses a more efficient algorithm that is better suited to the amount of rRNA data available today and in the foreseeable future. The new Probe Match accepts a candidate primer/probe, optionally with ambiguity codons, of up to 64 bases in length. While our previous version searched for hits within a specified number of mismatches (Hamming distance), the new version finds hits with a combination of mismatches and insertion/deletions (edit distance). Since some single insertion/deletion may be no more deleterious than a single mismatch, this new capability offers a significant improvement in the detection of potential cross-hybridization. In our previous implementation, the high percentage of partial sequences in the database limited the program's utility; it was difficult to determine if database entries failed to match simply because the sequence was incomplete in the target region. In this new version, the users can restrict analysis to database entries containing sequence data for the candidate probe target region of the rRNA molecule. (However, the search is not limited to this region of the molecule.) Similar to the other new programs, the results are displayed in an interactive version of the RDP Hierarchy. Each taxonomic rank lists the total number of sequences searched and the number matching within a user-specified edit distance. This maximum edit distance can be changed on the fly. For any hierarchy node, users can switch to a detail view listing the matching sequences. A third format is suitable for download and import into spreadsheet or other programs.

## RDP-II ACCESS AND CONTACT

The RDP-II data and analysis services can be found at <http://rdp.cme.msu.edu/>. The RDP's mission includes user support. The address for email support is [rdpstaff@msu.edu](mailto:rdpstaff@msu.edu). Telephone support is available (+1 517 432 4998). The RDP-II staff may also be contacted via fax (+1 517 353 8957 Attn:RDP) or regular mail.

## ACKNOWLEDGEMENTS

We thank several individuals for their past contributions: Robin Gutell (and his colleagues), Chuck Parker, Paul Saxman, Bonnie Maidak, Tim Lilburn, Niels Larsen, Tom Macke, Michael J. McCaughey, Ross Overbeek, Sakti Pramanik, Scott Dawson, Mitch L. Sogin, Gary Olsen and Carl Woese. The RDP is supported through US DOE OBER/NIH grant DE-FG02-99ER62848 and NSF grant DBI-0328255.

## REFERENCES

1. Cole, J.R., Chai, B., Marsh, T.L., Farris, R.J., Wang, Q., Kulam, S.A., Chandra, S., McGarrell, D.M., Schmidt, T.M., Garrity, G.M. *et al.* (2003) The Ribosomal Database Project (RDP-II): previewing a new autoaligner that allows regular updates and the new prokaryotic taxonomy. *Nucleic Acids Res.*, **31**, 442–443.
2. Brown, M.P.S. (2000) Small subunit ribosomal RNA modeling using stochastic context-free grammar. In: *Proceedings of the 8th International Conference on Intelligent Systems for Molecular Biology (ISMB 2000)*. San Diego, CA, pp. 57–66.

Table 1. rRNA search performance

Program <sup>a</sup>	Percentage of 16S rRNA queries <sup>b</sup> returning the most similar sequence <sup>c</sup> among the highest scoring <i>N</i> results		
	<i>N</i> = 1	<i>N</i> = 10	<i>N</i> = 20
Sequence Match	65	92	95
BLAST	39	53	55

<sup>a</sup>For both programs, the dataset consisted of 37 456 near full-length ( $\geq 1200$  base) rRNA sequences from the RDP release 9.20 alignment database.

<sup>b</sup>Query sequences (1000) were selected at random from the dataset.

<sup>c</sup>The most similar sequence to each query was determined by exhaustive pairwise similarity comparison of each query against the dataset. In cases of a tie in pairwise similarity, we required only one of the ties to be returned by the program.

3. Cannone, J.J., Subramanian, S., Schnare, M.N., Collett, J.R., D'Souza, L.M., Du, Y., Feng, B., Lin, N., Madabusi, L.V., Muller, K.M. *et al.* (2002) The Comparative RNA Web (CRW) Site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BioMed. Central Bioinformatics*, 3, 2.
4. Garrity, G.M., Bell, J.A. and Lilburn, T.G. (2004) *Taxonomic Outline of the Prokaryotes. Bergey's Manual of Systematic Bacteriology*, 2nd edn. Release 5.0, May 2004. Springer-Verlag, NY.
5. Wheeler, D.L., Chappey, C., Lash, A.E., Leipe, D.D., Madden, T.L., Schuler, G.D., Tatusova, T.A. and Rapp, B.A. (2000) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.*, 28, 10–14.
6. Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, 25, 3389–3402.

# Complete nucleotide sequence of a 16S ribosomal RNA gene from *Escherichia coli*

(recombinant plasmids/DNA sequence analysis/*rnmB* cistron)

JÜRGEN BROSIUS, MARGARET L. PALMER, POINDEXTER J. KENNEDY, AND HARRY F. NOLLER

Thimann Laboratories, University of California, Santa Cruz, California 95064

Communicated by Robert L. Sinshetmer, July 26, 1978

**ABSTRACT** The complete nucleotide sequence of the 16S rRNA gene from the *rnmB* cistron of *Escherichia coli* has been determined by using three rapid DNA sequencing methods. Nearly all of the structure has been confirmed by two to six independent sequence determinations on both DNA strands. The length of the 16S rRNA chain inferred from the DNA sequence is 1541 nucleotides, in close agreement with previous estimates. We note discrepancies between this sequence and the most recent version of it reported from direct RNA sequencing [Ehresmann, C., Stiegler, P., Carbon, P. & Ebel, J. P. (1977) *FEBS Lett.* 84, 337-341]. A few of these may be explained by heterogeneity among 16S rRNA sequences from different cistrons. No nucleotide sequences were found in the 16S rRNA gene that cannot be reconciled with RNase digestion products of mature 16S rRNA.

rRNA is becoming increasingly important in our current perception of the mechanism of action of ribosomes. In particular, we wish to understand more fully the functional role of 16S rRNA, the major molecular component of the small ribosomal subunit of prokaryotes. 16S rRNA has been directly implicated in discrimination of mRNA initiation sites (1, 2), tRNA binding (3, 4), and association of the two ribosomal subunits (5, 6). Full understanding of the workings of the ribosome consequently are becoming limited by our lack of knowledge of rRNA structure. Such information will also be essential for elucidation of the process by which these complex structures assemble themselves. Finally, we expect to gain insight into such diverse problems as protein-nucleic acid recognition and the evolutionary origin of the coding process from a more thorough knowledge of rRNA structure.

Partial nucleotide sequences for 16S RNA have been published (7-9). However, numerous discrepancies in oligonucleotide sequences (6, 10, 11) and in the ordering of oligonucleotides (6, 12-14) have been reported by other investigators. Additional evidence for sequence errors comes from restriction endonuclease mapping of a 16S RNA gene: some cleavage sites predicted from the published sequences were not found (unpublished data). Finally, chemical (6, 12, 15) and enzymatic (7, 8, 16) probes have been used to detect single-stranded regions of 16S RNA; the lack of agreement of these findings with the proposed secondary structure derived from the published sequences further suggests error in the primary structure. In view of the possible misinterpretation of various biochemical studies involving 16S RNA, we were prompted to reinvestigate its primary structure.

The availability of rapid DNA sequence methods made possible the derivation of the nucleotide sequence of 16S RNA by direct sequencing of the 16S rRNA gene from the *rnmB* cistron of *Escherichia coli*. We present here the complete sequence of the portion of the cistron corresponding to mature 16S RNA. Reported discrepancies with the previously published

sequence have been confirmed, and additional errors have been found involving oligonucleotide sequences, ordering of oligonucleotides, and, in one instance, the location of a larger section of the primary structure. No nucleotide sequences were found that cannot be accounted for from the RNase digestion products of mature 16S rRNA.

## METHODS

**Cloning and Mapping of DNA.** The 16S rRNA gene from the *rnmB* cistron of *E. coli* was cloned from two *EcoRI* restriction fragments of  $\lambda$ rf<sup>d</sup>18 (17, 18) in the ColE1 plasmid vector. Determination of the location of the 16S rRNA sequences and restriction enzyme cleavage sites will be described elsewhere. The small *HindIII* fragment from pER24 was excised and reinserted into the vehicle pBR322 (19) to give the recombinant plasmid pKK115 (see Fig. 1). All recombinant DNA experiments were carried out under P1, EK1 conditions, as specified in the National Institutes of Health guidelines.

**Generation of DNA Restriction Fragments.** Plasmid DNA was isolated and digested on a milligram scale with *EcoRI* or *HindIII*, and the inserted sequence was resolved from the cloning vector by sucrose gradient centrifugation (20). The purification of restriction enzymes *Alu* I, *Hpa* II, *EcoRI*, *Sma* I, *Bgl* II, and *HindIII* will be described elsewhere; *Hae* III, *Hha* I, *Hinf* I, and *Mbo* II were purchased from New England Biolabs; *Ava* II and *Taq* I were from Bethesda Research Laboratories. The cloned fragments (20-40 pmol) were digested with the appropriate restriction endonuclease to obtain DNA fragments of suitable size for sequence determination and isolated as described below.

**Sequencing by Primed Synthesis with DNA Polymerase.** For sequencing by either of the two methods developed by Sanger and coworkers (21, 22), restriction fragment primers were isolated by electrophoresis on 1% agarose gels (20 × 20 × 0.3 cm) in 0.02 M Na acetate/0.04 M Tris acetate, pH 8.3/2 mM EDTA or on 8% acrylamide/0.26% bisacrylamide gels (20 × 40 × 0.15 cm) in 0.09 M Tris borate, pH 8.3/2.5 mM EDTA. DNA bands were located by ethidium bromide staining, excised, crushed, and eluted by shaking overnight at room temperature in 0.5 M NaCl/0.1 M Tris-HCl, pH 8.0/5 mM EDTA and recovered by ethanol precipitation. Template strands were prepared from pER18 DNA by digestion with *Bgl* II, which cleaves at a single site in this plasmid, followed by cesium chloride density gradient centrifugation after heating and quick cooling in the presence of poly(U,G) (Miles) as described (23). Deoxynucleoside triphosphates labeled in the  $\alpha$  position with <sup>32</sup>P (200-300 Ci/mmol) were obtained from New England Nuclear. Dideoxynucleoside triphosphates were obtained from P-L Biochemicals; DNA polymerase I and its proteolytic fragment lacking 5' exonuclease activity were from Boehringer. Sequence determinations by the "plus and minus" method or by the "terminator" method were carried out as described (21, 22).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U. S. C. §1734 solely to indicate this fact.



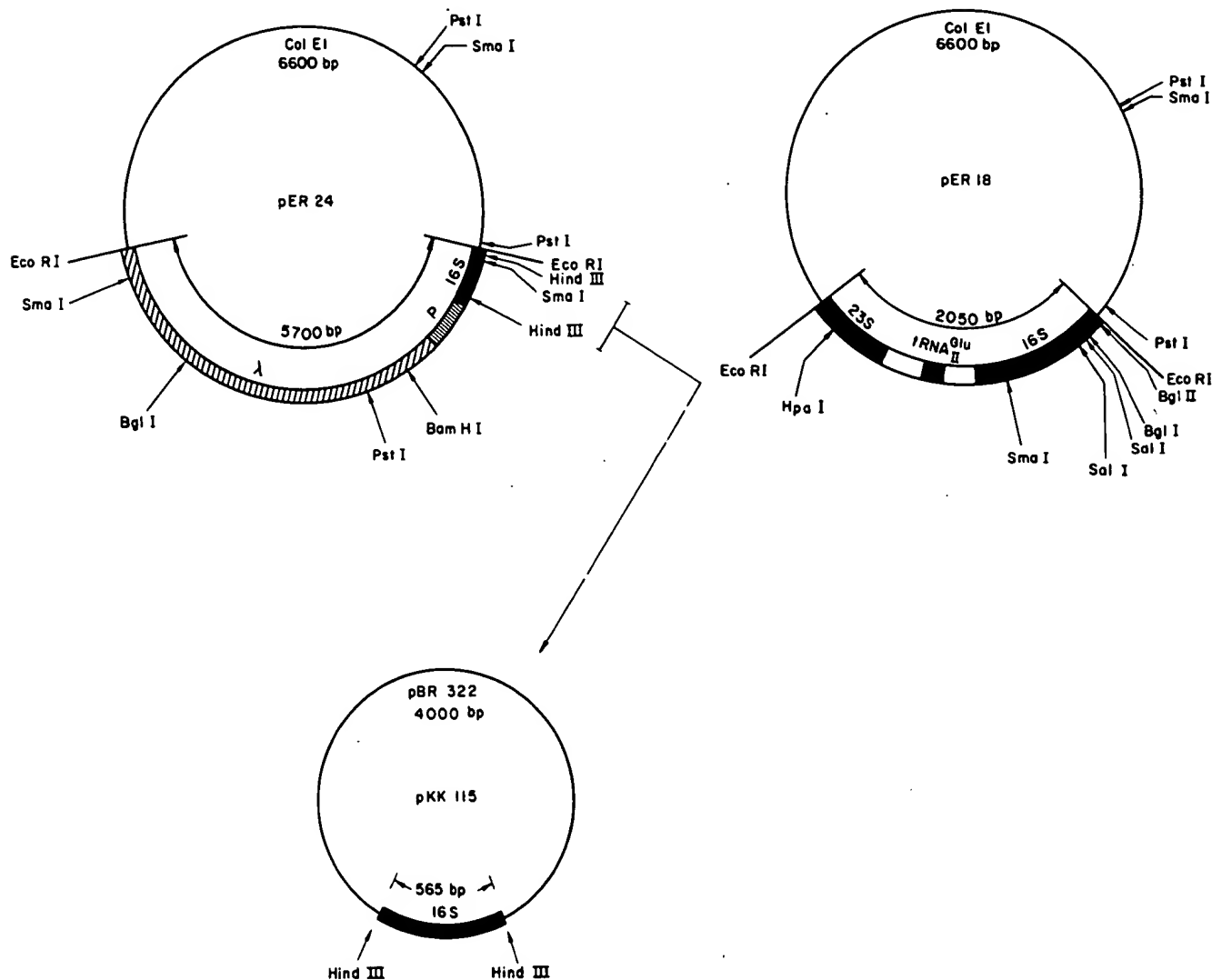


FIG. 1. Physical maps of recombinant plasmids containing portions of the 16S rRNA gene from the *rrnB* cistron of *E. coli*. The positions of 16S and 23S rRNA and tRNA sequences are shown by black segments. Sites of cleavage by restriction endonucleases are denoted by arrows. Wild-type  $\lambda$  sequences are shown by hatched segments, and the approximate position of the promoter for the *rrnB* cistron is indicated by the letter P. bp, base pairs.

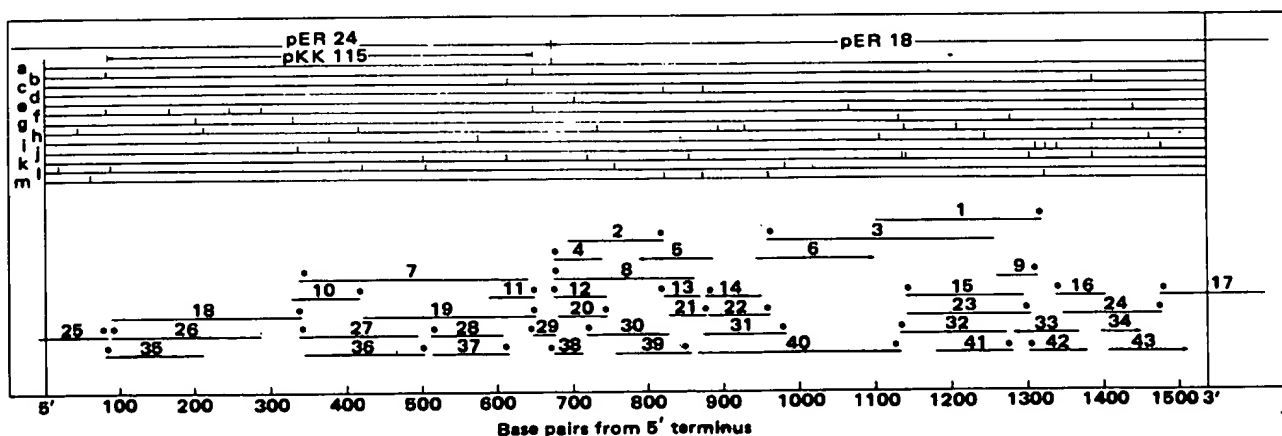


FIG. 2. Schematic map of the 16S rRNA gene, showing the detailed positions of restriction endonuclease cleavages, and the DNA fragments actually used in the sequence determination. Sequences obtained by primed synthesis (21, 22) are indicated by horizontal bars ending in arrows showing the direction of the chain extension. Sequences derived from partial chemical modification (24) are shown with asterisks at the 5' end of the labeled chain. Restriction endonucleases: a, *EcoRI*; b, *HindIII*; c, *Sma I*; d, *Sal I*; e, *Bgl II*; f, *Alu I*; g, *Ava II*; h, *Hae III*; i, *Hha I*; j, *HinfI*; k, *Hpa II*; l, *Mbo II*; m, *Taq I*. DNA fragments: 1, *Taq I*-3f; 2, *Taq I*-7f; 3, *Taq I*-3s; 4, *Alu I*-4f; 5, *Hae III*-13/L; 6, *Hae III*-13/H; 7, *HinfI*-1s; 8, *HinfI*-1s; 9, *HinfI*-1f; 10, *Hae III*-2f; 11, *Hha I*-3f; 12, *Mbo II*-12s; 13, *Taq I*-10s; 14, *Taq I*-9f; 15, *Hpa II*-4f; 16, *HinfI*-7f; 17, *HinfI*-3f; 18, *HinfI*-2f; 19, *Hae III*-1f; 20, *Mbo II*-12f; 21, *Taq I*-10f; 22, *Taq I*-9s; 23, *Hpa II*-4s; 24, *HinfI*-7s; 25, *HindIII/BamHI*-2; 26, *HinfI*-2s; 27, *HinfI/Hpa II*-2; 28, *Hpa II*-2f; 29, *HindIII/BamHI*-1; 30, *Hpa II*-6s; 31, *Mbo II*-5s; 32, *Ava II*-4s; 33, *Sma I/EcoRI*-1/L; 34, *Sma I/EcoRI*-2/H; 35, *Hae III*-3f; 36, *Hpa II/HinfI*-2; 37, *Hpa II*-2s; 38, *Hpa II*-13f; 39, *Hpa II*-6f; 40, *Ava II/Hpa II*-2; 41, *Ava II*-4f; 42, *Hpa II*-9f; 43, *Sma I/EcoRI*-2/H. The portions of the sequence covered by the plasmid DNA molecules in Fig. 1 are also shown.

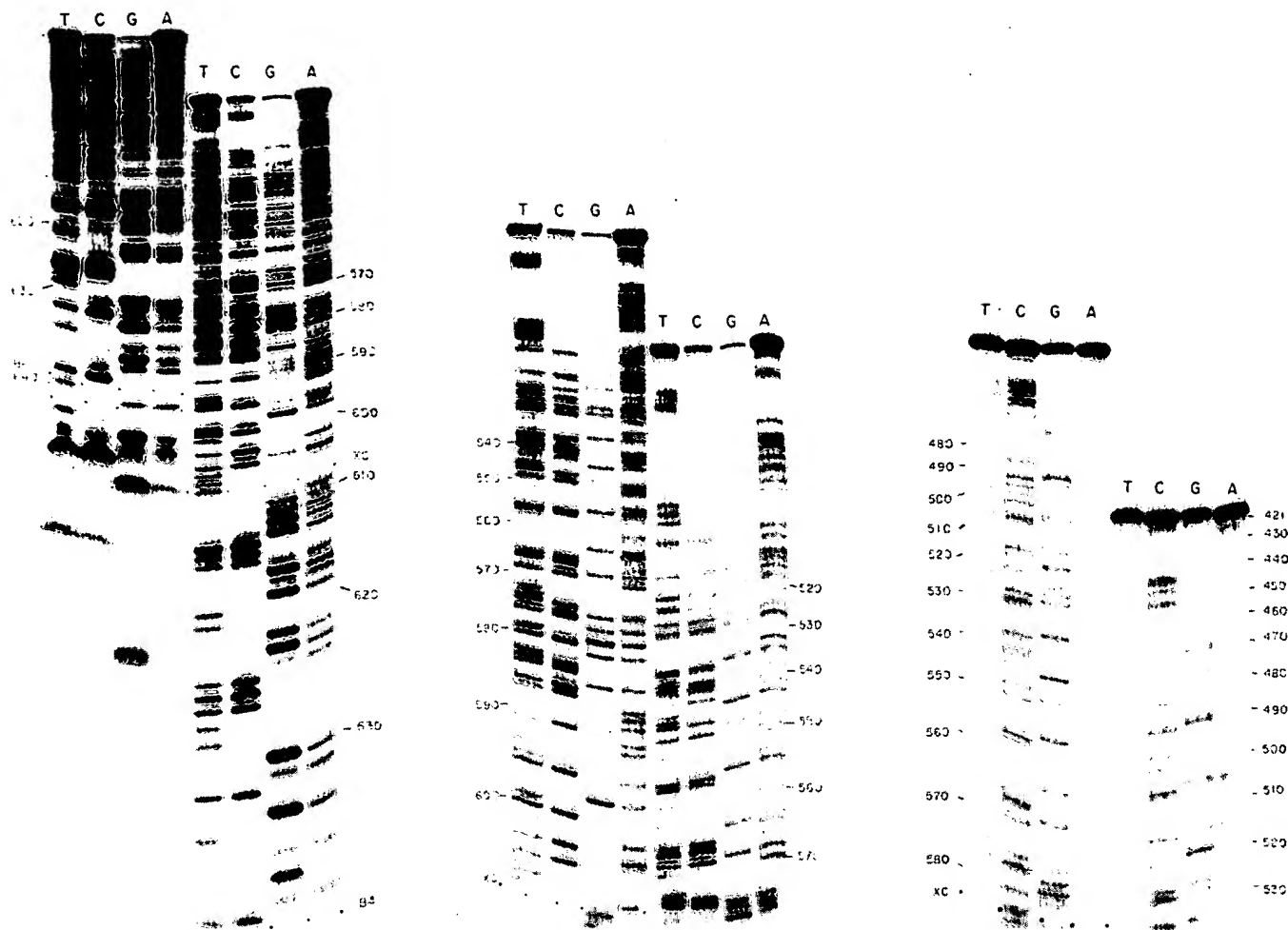


FIG. 3. Autoradiographs of sequence gels from the partial chemical modification method applied to the single-stranded DNA fragment *Hae* III-1f (Fig. 2). The sequence read from the gel is complementary to the 16S RNA sequence and covers the region from positions 421 to 650. The positions of nucleotides in the 16S RNA sequence are indicated by the numbers. The gels ( $0.7 \times 20 \times 40$  cm) were 12% polyacrylamide/0.6% bisacrylamide run 1, 3, 9, and 12 hr, respectively, at 1000 V (Left and Center) and 8% acrylamide/0.4% bisacrylamide run for 7 and 10 hr, respectively, at 1000 V (Right).

**Sequencing by the Partial Chemical Modification Method.** For sequence determination according to the method of Maxam and Gilbert (24), the mixture of restriction fragments (20–40 pmol) was treated with bacterial alkaline phosphatase (1  $\mu$ g/5  $\mu$ g of DNA) for 1 hr at 37°, extracted four times with an equal volume of phenol saturated with 10 mM Tris-HCl, pH 8.0/1 mM EDTA, and precipitated with ethanol. The fragments were then labeled at their 5' termini by using 5–10 units of polynucleotide kinase (P-L Biochemicals) and 0.5–2 mCi of [ $^{32}$ P]ATP (1500–2000 Ci/mmol) synthesized as described (25). Excess ATP and phosphate were removed by passing the reaction mixture through a column of Sephadex G-75 in 2.5 mM Tris-HCl/1 mM NaCl, pH 7.4, poured in a 1-ml disposable pipet. The peak containing the labeled DNA was lyophilized and taken up in 50  $\mu$ l of 0.02 M Tris borate, pH 8.3/0.5 mM EDTA, and the labeled restriction fragments were resolved by electrophoresis on an 8% polyacrylamide gel ( $0.5 \times 20 \times 40$  cm) as described above. DNA fragments were located by radioautography and eluted as described above. Singly end-labeled DNA was obtained by strand separation on 8% acrylamide/0.13% bisacrylamide gels ( $0.5 \times 20 \times 40$  cm) (26) after the DNA sample was dissolved in 0.3 M NaOH/1 mM EDTA/10% (vol/vol) glycerol/0.05% xylene cyanol/0.05% bromophenol blue or by recleavage of the double-stranded DNA with a second restriction endonuclease.

Chemical treatment of the end-labeled DNA was performed

essentially as described (24) except that a saturated solution of NaCl was used for cytidine suppression, 1 M instead of 0.5 M piperidine was used for the pyrimidine reactions, and piperidine was removed under reduced pressure at 60° instead of by lyophilization. Sequence gels were constructed and run essentially as described (27).

## RESULTS AND DISCUSSION

**Sequencing Strategy.** The 16S RNA gene from the *E. coli* *rnmB* cistron was conveniently isolated in two *Eco*RI fragments from the transducing phage  $\lambda$ rif<sup>d</sup>18 (17, 18). These fragments were cloned by insertion at the *Eco*RI site of ColE1 plasmid, as shown in Fig. 1. Because of the difficulty in growing large amounts of the recombinant plasmid pER24, presumably due to the presence of phage genes in the plasmid, the small *Hind*III fragment was excised from pER24 and reinserted into the vehicle pBR322 (19). Residues 1–80 and 648–673 and the sequence preceding the 5' terminus of 16S RNA were obtained from pER24, residues 81–647 were from pKK115, and residues 674–1541 and the sequence following the 3' terminus were from pER18. Fig. 2 summarizes the positions of cleavage by the restriction endonucleases used in this work and the fragments from which the sequences were derived.

Most of the sequence was derived by the partial chemical modification method of Maxam and Gilbert (24). The sequence was determined independently two to six times at each position,

<sup>L</sup>AAUUGAAGAGUUGAUCAU<sup>U</sup>GGCUCAGAU<sup>U</sup>GAACGCUGG<sup>C</sup>GGCAGGCCUA 50  
 ACACAUGCAAGUCGAACGGUAACAGGAAGAAGCUUG<sup>U</sup>CUUUUGCUGACGA 100  
 GUGGCGGACGGGUGAG<sup>U</sup>AAUGUCUGGGAAACUGCCUGAUGGAAGGGG<sup>U</sup>AA 150  
 ACUACUGGAAACGGUAGCUAAUACCG<sup>C</sup>CAUAAACGUCGCAAGACC<sup>A</sup>AAAGAGG 200  
 GGGACC<sup>U</sup>UCGGGGCCUCUUGCCAUCGGAUGUG<sup>C</sup>CCAGAUUGGGAUUAGCUAG 250  
 UAGGUGGGGUAACG<sup>G</sup>CUACCUAGGCGACGAUCCCU<sup>G</sup>AGCUGGUCUGAGAG 300  
 G<sup>A</sup>UGACCAGCCACACUGGAACUGAGACACGGUCCAGACUCCUACGGGAGG 350  
 CAGCAGUGGG<sup>G</sup>GAUAU<sup>U</sup>UGCACAAUGGGCGCAAGCCUGAUGCAGCCAUGCC 400  
 GCGUG<sup>U</sup>AUGAAGAAGGCCUUCGGGUUGUAAAG<sup>U</sup>ACU<sup>U</sup>UACGCGGGGAGG<sup>A</sup> 450  
 AGGGAGUAAA<sup>G</sup>GUUAAU<sup>A</sup>CCUUGCUCAUUGACGUUACCCGCAGAAAGAGC 500  
 ACCGG<sup>C</sup>UAACUCCGUGCCAGCAGC<sup>C</sup>CGGUAAUACGGAAGGGUGCAAGCGU 550  
 UAAUCGGAAUUAUCUGGGCGUAAAGCGCACGCAGG<sup>C</sup>GGUUGUUAAGUCAG 600  
 AUGUGAAAUCCCCGGGCUCAACCUGGGAACUGCAUCUGAUACUGGCAAGC 650  
 UUG<sup>A</sup>GUUCUCGUAGAGGGGGGUAGAAUCCAGGUGUAGCGGUGAAAU<sup>G</sup>CG<sup>U</sup> 700  
 ABAGAUCUGGAGGAAUACCG<sup>G</sup>UGGCGAAGGCGGCCCCUGGACGAAGACU 750  
 GACGCUCAGGUGCGAAAGCGUGGGGAG<sup>C</sup>AAACAGGAUUAG<sup>A</sup>UACCCUGGU 800  
 AGUCCACGCCGUAACGAUGUCGACUUGGAGGUUGUGCCCUUG<sup>A</sup>GGCGUG 850  
 GCUUCCGGAGCUAAACCGGUUAAAGUCGACCGCCUGGGGAGUACGGCCGCAA 900  
 GGUUAAAACUCUAAAU<sup>G</sup>AAUUGACGGGGGCCCGCACAAAGCGGUGGAGCAUG 950  
 UGGUUAUUAUUCGAU<sup>U</sup>AG<sup>C</sup>AAACGCGAAGAACC<sup>U</sup>UACCGGUUUG<sup>A</sup>CAUCCAC 1000  
 GGAAGUUUUCAGAGAUGAGAAUGUGCCUUCGGGAACCGUGAGACAG<sup>G</sup>UGC 1050  
 UGCAUGGCUGUGCUCAGCUCGUGUUGUG<sup>A</sup>AAUGUUGGGUUAAGUCCCGCA 1100  
 ACGAGCGCAACCCUUAUCCUUGUUGCCAGCGGUCGGCCGGGAACUCAA 1150  
 AGGAGACUGCCAGUG<sup>A</sup>UAAACUGGAGGAAGGUGGGGAUGACGUCAAGUCA 1200  
 UCAUG<sup>C</sup>CCC<sup>U</sup>UACGACCGGGCUACACACGUGCUACAAGGCGCAUACAA 1250  
 AGAGAAGCGACCUCGCG<sup>A</sup>GAGCAAGCGG<sup>A</sup>CCUCAUAAAGUGCGUCGUAGU 1300  
 CCGGAUUGGAGUCUGCAACUCGACUCCAUGAAGUCGGAUUCGUAGUAAU 1350  
 CGUGGAUCAG<sup>A</sup>AUGCCACGGUGAAUACGUUCCCGGGCCUUGUACACACCG 1400  
<sup>A</sup>CCGUG<sup>C</sup>ACACCAUGGGAGUGGGUUGCAAAGAAGUAGGUAGCUUAACCUU 1450  
 CGGGAGGGCGCUUACCACUUGUGAUUCAUGACUGGGGGUGAAG<sup>U</sup>CGUAAC 1500  
 AAGGUAAACCGUAGGGG<sup>G</sup>GA<sup>A</sup>CCUGCGGUUGGAUACCCUCCUUA<sup>OH</sup> 1541

FIG. 4. The nucleotide sequence of 16S rRNA from the *rrnB* cistron of *E. coli*. The RNA sequence was inferred from the DNA sequence. The ends of the 16S rRNA and positions of methylated nucleotides were identified by comparison with RNA sequence results (1, 6–10, 28–31, and unpublished). For comparison with the sequences reported by Ehresmann *et al.* (7–9), the positions of their lettered sections is given. In some cases (e.g., section I<sup>U</sup>) the assignment of these sections was necessarily approximate.

and by sequencing both DNA strands, except for residues 1–80, 288–336, 652–673, and 1513–1541, which were sequenced on only one strand.

**Sequence Results.** Examples of sequence gels are shown in Fig. 3. By use of thin gels (27) and other minor modifications, it was often possible to read sequences accurately to about 250 base pairs from the proximal end of a fragment. Fig. 4 shows the complete nucleotide sequence of the 16S RNA as derived from the gene sequence. The total length of the chain is 1541 nucleotides, in close agreement with earlier estimates based on chemical methods (9, 32) but somewhat less than values predicted from physicochemical methods (33–35).

**Comparison with Previous Sequence Results.** Two conflicting catalogs of RNase T1 oligonucleotides have been published for *E. coli* 16S RNA (8, 36). In the course of determining sequences around kethoxal-reactive sites in 16S RNA (6, 12), we have noted some discrepancies between our results and those of Ehresmann *et al.* (8), but our results are in agreement with

the sequences reported by Woese and coworkers (11, 36). In the most recently published version of their sequences, Ehresmann *et al.* (9) reported much closer agreement with the latter catalog, but a few discrepancies remain. There are no differences between the final catalog of T1 oligonucleotides as determined by the Woese group and those in the present sequence (Fig. 4), with the apparent exception of two modified oligonucleotides, whose general forms are C-C-N-C-G (positions 524–528) and N-C-C-G (positions 1401–1404) (refs. 10 and 11; C. Woese, personal communication). The differences, however, are trivial in that they concern only the identification of the modified nucleotides itself. Woese and coworkers incorrectly assumed that N in the pentanucleotide corresponds to <sup>4</sup>mC<sub>m</sub> and therefore N in the tetranucleotide is <sup>7</sup>mG (C. Woese, personal communication).

A serious source of discrepancy between the sequence presented here and that by Ehresmann *et al.* (9) is in the alignment of oligonucleotides. This has been previously noted in instances

in which kethoxal modification studies have provided ordering of adjacent oligonucleotides (6), in a recent study using rapid RNA sequencing techniques (13), and in other studies using DNA methods (14). A prominent example of this problem is the region 398–504 (Fig. 4). In the latest version of the sequence of Ehresmann *et al.* (9), 43 nucleotides originating from positions 462–504 are inserted at position 432/433.

The remaining differences between the sequence presented here and the previous one (9) concern small sections of the molecule, and many involve the insertion or deletion of mono- or dinucleotides. In all, we note discrepancies involving about 200 nucleotides. A few of these differences might be due to sequence heterogeneity among the various rRNA cistrons, for which evidence has been presented (7–10, 34). Our sequence is in agreement with much of the recently published corrections in the region 1103–1165 and 1414–1488 reported by Ross and Brimacombe (13) on the basis of a newly developed rapid RNA sequence method. However, these authors have apparently deleted pyrimidines at positions 1161, 1466, 1473, 1477, 1478, and 1480. Young and Steitz (14) used DNA sequence methods and reported sequences from the 5' and 3' ends of 16S RNA from *rnnD* and one other rRNA cistron, as yet unidentified. Their results for residues 1–18 and 1381–1541 are in complete agreement with our findings.

**Consequences of Sequence Changes for Secondary Structure Prediction.** Previously, we noted that a majority of the sites of kethoxal modification of 16S RNA, which must be single stranded (37), have been assigned to double-stranded structures in the secondary structure model for 16S RNA proposed by Ehresmann *et al.* (8). Furthermore, some of the sites of nuclease attack observed under partial digestion conditions reported by these authors (7, 8) are inconsistent with their model. As anticipated, the sequence in Fig. 4 gives rise to a very different secondary structure prediction, in which these conflicts are largely resolved. A secondary structure model based on this sequence as well as a discussion of the structural and functional implications of these findings will be presented elsewhere.

**Possible Coding Functions of 16S RNA.** The use of rRNA as messenger by the cell has often been considered. Examination of the 16S rRNA sequence in Fig. 4 shows that most AUG or GUG triplets are followed by in-phase termination codons after a short interval. One exception is the AUG triplet at position 1187, which has no in-phase terminator until position 1439, allowing for a translation of a sequence of 84 amino acids. Furthermore, preceding the initiator codon by a distance of nine nucleotides is the sequence G-G-A-G-G, which would allow a favorable base-paired interaction with the 3'-terminal sequence of 16S rRNA in the mRNA recognition mechanism proposed by Shine and Dalgarno (1). Whether this theoretically possible message is in fact translatable *in vivo* is open to question. We have searched the sequence for possible homology with known *E. coli* protein sequences, including the ribosomal proteins, without success. It is most likely that translation of this sequence is precluded by formation of stable secondary structure of the RNA and by binding of ribosomal proteins.

We are grateful to the following friends and colleagues for supplying unpublished results, strains, technical advice, and encouragement at various stages of this work: J. Abelson, W. Barnes, B. Barrell, W. Gilbert, C. Guthrie, W. Herr, K. Katze, J. Kirschbaum, B. Lowe, A. Maxam, A. Otsuka, C. Panofsky, R. Rodriguez, F. Sanger, S. Stein, J. Steitz, C. Woese, and R. Young. J.B. is a postdoctoral fellow of the Fogarty International Center of the National Institutes of Health; M.L.P. is the

recipient of a President's Undergraduate Fellowship from the University of California. This work was supported by National Institutes of Health Grant GMS-17129.

- Shine, J. & Dalgarno, L. (1974) *Proc. Natl. Acad. Sci. USA* 71, 1342–1346.
- Steitz, J. A. & Jakes, K. (1975) *Proc. Natl. Acad. Sci. USA* 72, 4734–4738.
- Noller, H. F. & Chaires, J. B. (1972) *Proc. Natl. Acad. Sci. USA* 69, 3115–3118.
- Schwartz, I. & Ofengand, J. (1974) *Proc. Natl. Acad. Sci. USA* 71, 3951–3955.
- Santer, M. & Shane, S. (1977) *J. Bacteriol.* 130, 900–910.
- Chapman, N. M. & Noller, H. F. (1977) *J. Mol. Biol.* 109, 131–149.
- Ehresmann, C., Stiegler, P., Fellner, P. & Ebel, J. P. (1972) *Biochimie* 54, 901–967.
- Ehresmann, C., Stiegler, P., Fellner, P., & Ebel, J. P. (1975) *Biochimie* 57, 711–748.
- Ehresmann, C., Stiegler, P., Carbon, P. & Ebel, J. P. (1977) *FEBS Lett.* 84, 337–341.
- Uchida, T., Bonen, L., Schaup, H. W., Lewis, B. J., Zablen, L. & Woese, C. (1974) *J. Mol. Evol.* 3, 63–77.
- Magrum, L., Zablen, L., Stahl, D. & Woese, C. (1975) *Nature (London)* 257, 423–426.
- Noller, H. F. (1974) *Biochemistry* 13, 4694–4703.
- Ross, A. & Brimacombe, R. (1978) *Nucleic Acids Res.* 5, 241–256.
- Young, R. A., & Steitz, J. A. (1978) *Proc. Natl. Acad. Sci. USA* 75, 3593–3597.
- Hogan, J. J. & Noller, H. F. (1978) *Biochemistry* 17, 587–593.
- Santer, M. & Santer, U. (1973) *J. Bacteriol.* 117, 63–77.
- Kirschbaum, J. B. & Konrad, E. B. (1973) *J. Bacteriol.* 116, 517–526.
- Lindahl, L., Jaskunas, S. R., Dennis, P. P. & Nomura, M. (1975) *Proc. Natl. Acad. Sci. USA* 72, 2743–2747.
- Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. C., Heyneker, H. L. & Boyer, H. W. (1977) *Gene* 2, 95–113.
- Valenzuela, P., Bell, G. I., Venegas, A., Sewell, E. T., Masiaz, F. R., DeGennaro, L. J., Weinberg, F. & Rutter, W. J. (1977) *J. Biol. Chem.* 252, 8126–8135.
- Sanger, F. & Coulson, A. R. (1975) *J. Mol. Biol.* 94, 441–448.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* 74, 5463–5467.
- Szybalski, W. (1968) *Methods Enzymol.* 12, 330–360.
- Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* 74, 560–564.
- Glynn, I. M. & Chappell, J. B. (1964) *Biochem. J.* 90, 147–149.
- Szalay, A. A., Grohmann, K. & Sinsheimer, R. L. (1977) *Nucleic Acids Res.* 4, 1569–1578.
- Sanger, F. & Coulson, A. R. (1978) *FEBS Lett.* 87, 107–110.
- Noller, H. F. & Herr, W. (1974) *Mol. Biol. Rep.* 1, 437–439.
- Ehresmann, C., Stiegler, P. & Ebel, J. P. (1974) *FEBS Lett.* 49, 47–48.
- Fellner, P. (1969) *Eur. J. Biochem.* 11, 12–27.
- Sprague, K. U. & Steitz, J. A. (1975) *Nucleic Acids Res.* 2, 787–798.
- Midgley, J. E. M. (1965) *Biochim. Biophys. Acta* 108, 340–347.
- Kurland, C. G. (1960) *J. Mol. Biol.* 2, 83–91.
- Stanley, W. M. & Bock, R. M. (1965) *Biochemistry* 4, 1302–1311.
- Ortega, J. P. & Hill, W. E. (1973) *Biochemistry* 12, 3241–3243.
- Woese, C. R., Fox, G. E., Zablen, L., Uchida, T., Bonen, L., Pechman, K., Lewis, B. J. & Stahl, D. (1975) *Nature (London)* 254, 83–86.
- Litt, M. (1969) *Biochemistry* 8, 3249–3253.

## Bacterial Evolution

CARL R. WOESE

*Department of Microbiology, University of Illinois, Urbana, Illinois 61801*

<b>PERSPECTIVE</b> .....	222
A Fruitless Search and Its Consequences.....	222
Three Ideas That Shape Our Concept of Bacterial Evolution.....	224
Procaryote-eucaryote dichotomy .....	224
Oparin Ocean scenario .....	224
Darwin's warm little pond .....	226
<b>MEASUREMENT OF BACTERIAL PHYLOGENETIC RELATIONSHIPS</b> .....	226
Three-Dimensional and One-Dimensional Characters .....	226
The evolutionary clock .....	226
A missed opportunity .....	226
Nature of Molecular Chronometers .....	227
Clocklike behavior.....	227
Phylogenetic range .....	227
Size and accuracy.....	227
rRNAs, the Ultimate Molecular Chronometers .....	227
Why they are so good .....	227
Oligonucleotide cataloging.....	228
Analysis of Sequence Alignments .....	228
Distance matrix methods.....	228
Maximum parsimony analysis .....	228
Cluster analysis .....	228
<b>DO BACTERIA HAVE A GENEALOGY AND A MEANINGFUL TAXONOMY?</b> .....	230
<b>THE UNIVERSAL PHYLOGENETIC TREE</b> .....	231
<b>EUBACTERIAL PHYLOGENY</b> .....	232
Background .....	232
Major Eubacterial Groups and Their Subdivisions .....	232
Purple Bacteria.....	236
$\alpha$ -Purple bacteria .....	237
$\beta$ -Purple bacteria .....	238
$\gamma$ -Purple bacteria .....	238
$\delta$ -Purple bacteria .....	238
Gram-Positive Eubacteria .....	240
Cyanobacteria .....	242
Green Sulfur Bacteria .....	243
Spirochetes .....	244
Bacteroides, Flavobacteria, and Their Relatives.....	244
Planctomyces and Their Relatives.....	244
Chlamydiae.....	245
Radiation-Resistant Micrococci and Their Relatives.....	245
Green Non-Sulfur Bacteria and Their Relatives .....	245
Other Eubacterial Phyla .....	245
Overall Structure of the Eubacterial Tree.....	245
<b>ARCHAEBACTERIAL PHYLOGENY</b> .....	247
Unusual Nature of the Archaeobacterial Phenotype.....	247
Definition of the Major Archaeobacterial Groups .....	247
Methanogens .....	248
Extreme halophiles .....	248
Extreme thermophiles .....	249
The Archaeobacterial Tree: Its Branching Order and Root.....	249
Main branches.....	249
<i>Thermococcus</i> and <i>Thermoplasma</i> .....	251
Branching order among the methanogens and extreme halophiles.....	251
<b>A BRIEF LOOK AT EUCARYOTE PHYLOGENY</b> .....	252
Implications of the Eucaryote Phylogenetic Tree .....	252
<b>NATURE OF THE EVOLUTIONARY PROCESS IN BACTERIA</b> .....	253
Relationship Between Evolution's Tempo and Its Mode.....	253
Macroevolution at the Molecular Level: the Mycoplasmas and Their rRNAs.....	253

Dynamics of Variation in rRNA .....	255
Why mycoplasma rRNAs are so unusual .....	256
Summary .....	256
Both Tempo and Mode of Bacterial Evolution Are Reflected in rRNA .....	256
Macroevolution in mycoplasmas is chronic .....	256
Conditions for true macroevolution among bacteria .....	257
Basic Principles of Bacterial Evolution .....	257
Implications for Bacterial Taxonomy .....	257
EVOLUTION OF THE TWO PROCARYOTIC PHENOTYPES .....	258
Archaeobacterial Evolution .....	258
Ancestral phenotype .....	258
Evolution of methanogenic and halophilic phenotypes .....	258
Archaeobacterial ribosomes and their evolutionary implications .....	259
Summary .....	259
Eubacterial Evolution .....	259
Photosynthesis .....	259
Thermophilla .....	260
Conclusion .....	260
Differences between Archaeobacteria and Eubacteria .....	260
Cytological and physiological differences .....	260
Niches .....	261
Metabolic versatility .....	261
Molecular plasticity .....	261
Rates of evolution .....	261
Summary .....	261
Why Are Archaeobacteria and Eubacteria So Different? .....	261
Primitiveness of Archaeobacteria .....	262
When Did Procaryotes Evolve? .....	262
THE UNIVERSAL ANCESTOR .....	262
Progenotes .....	263
Characteristics of the progenote .....	263
Was the Universal Ancestor a Progenote? .....	263
ACKNOWLEDGMENTS .....	264
LITERATURE CITED .....	264

## PERSPECTIVE

A revolution is occurring in biology: perhaps it is better characterized as a revolution within a revolution. I am, of course, referring to the impact that the increasingly rapid capacity to sequence nucleic acids is having on a science that has already been radically transformed by molecular approaches and concepts. While the impact is currently greatest in genetics and applied areas such as medicine and biotechnology, its most profound and lasting effect will be on our perception of evolution and its relationship to the rest of biology. The cell is basically an historical document, and gaining the capacity to read it (by the sequencing of genes) cannot but drastically alter the way we look at all of biology. No discipline within biology will be more changed by this revolution than microbiology, for until the advent of molecular sequencing, bacterial evolution was not a subject that could be approached experimentally.

With any novel scientific departure it is important to understand the historical setting in which it arises—the paradigm it will change. Old prejudices tend to inhibit, distort, or otherwise shape new ideas, and historical analysis helps to eliminate much of the negative impact of the status quo. Such analysis is particularly important in the present instance since microbiologists do not deal with evolutionary considerations as a matter of course and so tend not to appreciate them. Therefore, I begin this discussion with a

brief look at how the relationship between microbiology and evolution (i.e., the lack thereof) developed.

## A Fruitless Search and Its Consequences

Microbiologists of the late 1800s and early 1900s were certainly as cognizant of evolutionary considerations as any biologists. They assigned as much importance to determining the natural (evolutionary) relationships among bacteria as zoologists and botanists did to determining metazoan genealogies. From Beijerinck to Kluver to van Niel, a main concern of the Dutch school, perhaps the dominant force in microbiology in the first half of this century, had been these natural relationships. And it must have been the hope of someday knowing them that inspired the founders of *Bergey's Manual* to adopt for bacteria the same classification system used to group animals and plants phylogenetically.

The search for a bacterial phylogeny, however, became mired in failure, generations of it. Animals and plants are rich in complex morphological detail, which served as the basis for their phylogenetic classification. Bacteria, on the other hand, have such simple morphologies that these are of no use in defining their phylogeny. Bacterial physiologies are more useful in this regard, but are still too limited; while shared physiological traits often correctly group bacterial species, relatives lacking the trait in question usually exist as

well (56). Some of the early microbiologists, realizing the pitfalls in attempting to classify bacteria by the then available criteria, avoided the area, which served to deemphasize the role of evolutionary considerations in the development of microbiology. Those who did concern themselves with phylogenetic classification of bacteria created distorted, basically flawed schemes which confused rather than resolved problems and ultimately discredited the whole attempt. (One of the delightful absurdities to emerge from this involuted system is the taxon *Pseudomonas*, perhaps the best known, most studied, and most pedagogically utilized "representative genus," which actually is a collection of at least five separate groups of bacteria [53], whose name derives from the Greek *pseudes* and *monas*, i.e., false unit!)

The situation seems to have reached a watershed during the time of C. B. van Niel. Initially a leader in attempts to determine microbial phylogenies (210), van Niel, in apparent frustration, ultimately gave up on them, settling for a determinative classification system the basic purpose of which was to identify species (226). Van Niel was perhaps the last microbiologist to treat the matter of bacterial evolution seriously.

Without a capacity to determine bacterial genealogies, considerations of bacterial evolution are mere cerebral exercises. Realizing this, students found it difficult to become enthusiastic about the subject. There was no way microbiology textbooks could treat it seriously either. Many did not treat it at all! Moreover, the failure to determine evolutionary relationships seemed to generate the feeling that it was not important to do so. Bacterial evolution was all but forgotten. All that remained to represent this once vital and important area of microbiology was a formal and unappealing bacterial taxonomy: one falsely authoritative in its borrowed use of the Linnaean (phylogenetic) classification system, stultifying in its liturgy, and caught up in classification for classification's sake.

The result of all this was that microbiology worked from a paradigm that for all intents and purposes was devoid of evolutionary concepts. They played no role in its pedagogy; they had no influence on the design and interpretation of experiments; they were not a part of its value structure. Roger Stanier, one of the few microbiologists who maintained any interest at all in bacterial evolution, captured the spirit of the times with this piquant proscription (written about 1970): "Evolutionary speculation constitutes a kind of metascience, which has the same intellectual fascination for some biologists that metaphysical speculation possessed for some mediaeval scholastics. It can be considered a relatively harmless habit, like eating peanuts, unless it assumes the form of an obsession; then it becomes a vice" (209). That microbiology had reduced evolutionary matters to the status of dalliance was indeed unfortunate, for much of what is important and interesting about evolution lay hidden in the microbial world.

Fortunately, nucleic acid sequencing technology today makes bacterial phylogeny a tractable problem. In fact, all phylogenetic relationships can now be determined much more easily and in far more detail and depth than was ever dreamed possible (116, 174). Microbiology is consequently being inundated with sequence information, which accumulates so rapidly that the reading and entering of data are becoming major concerns, while the actual sequencing operations will soon cease to be rate-limiting factors. The data are in a form [pages of A's, C's, G's, and T(U)'s] alien to most microbiologists, their analysis is arcane, and phylogenetic conclusions tend to be presented in a take-it-

or-leave-it manner. It is understandable that some microbiologists distrust these conclusions. However, it is not permissible to ignore them. Phylogenies derived from sequence analysis have to be accepted for what they minimally are: hypotheses, to be tested and either strengthened or rejected on the basis of other kinds of data (97).

Microbiology is now at a point at which it has to ask why it should concern itself with bacterial evolution and what such concern would mean to its future. The emerging bacterial phylogeny cannot be viewed as having merely local impact, i.e., as being a revision of existing bacterial taxonomy. At the very least the existing taxonomy will be totally rewritten by what is currently happening. Even that, however, is only the tip of the iceberg. Phylogenetic perspective will affect the microbiology paradigm throughout. This is already apparent in the change in perspective accompanying the discovery of archaeobacteria. Bacteria will no longer be conceptualized mainly in terms of their morphologies and biochemistries; their relationships to other bacteria will be central to the concept as well. Design and interpretation of experiments will be significantly changed. Microbial biochemistry will be conceptualized more in a comparative way. Medical microbiology will have a broadened perspective. Phylogenetic considerations will increase the microbiologist's interest in microbial ecology and shape his approach to it (157, 164). Perhaps the most significant change will be the altering of our perception of the relationship between procaryotes and eucaryotes and, therefore, of the position microbiology holds in relation to the other biological disciplines.

The evolutionist, too, needs to concern himself with the effect that opening the "Pandora's box" of bacterial evolution is going to have, for although he is exquisitely aware of evolution as it is encountered among the metazoa, the world of forms and fossils, bacterial evolution is as alien to him as it is to the microbiologist. Determining microbial phylogeny is not simply the long awaited completion of the "Darwinian programme," the extension of evolutionary study to all life on this planet. Rather than its providing the few missing pieces in the great puzzle of evolution, bacterial evolution in effect is the puzzle. It increases the current time span of evolutionary study by almost an order of magnitude. It holds the key to the origin of the eucaryotic cell. It shows the evolutionist an intimacy between the evolution of the planet and the life forms thereon that he has never before experienced and which, consequently, will lead to a close relationship between the geologist (not merely the paleontologist) and the evolutionist. It will redefine the classical question concerning the connection between evolutionary rate and the quality of the resulting change (the so-called tempo-mode problem) in new and more powerful terms (see below). Even the way we conceptualize selection, the roles of positive versus negative selection, may be changed. In other words, bacterial evolution will show us that, far from approaching the culmination of evolutionary study, where one refines existing concepts and fills in the details, we are only at its beginning.

The impact of the study of bacterial evolution should be felt throughout all of biology, bringing about major shifts in emphasis. The nature of the universal ancestor (which, if it is given attention at all today, is seen as just another common ancestor of a group) will be recognized for the important biological problem that it is. A major effort will be mounted to define the ancestral gene families. And the historical dimension will become a significant, useful part of the study of macromolecular structure. Because its conceptual base

now rests outside of biology per se (in physics and chemistry), biology's interests, its thrust, have tended increasingly to be defined by external factors, many even extrascientific. Biology has become very much an "other directed" discipline. What a renewed and broadened interest in evolutionary questions can and will do is to restore to biology an internally defined sense of direction.

Some restructuring at a metaphysical level is even a possibility. Biology's base, scientific materialism (that highly reductionist, highly mechanistic picture borrowed wholesale from the 19th century physicist), is a world view to which physicists since Einstein, Bohr, and Schrödinger can no longer subscribe. If there by anything to Whitehead's concept of evolution as basic process (236), the interest in evolution generated by its study in bacteria (and unicellular eucaryotes) may push biology in this direction, toward a process-oriented outlook, an attitude that processes (evolution, development, mind) somehow underlie genes, cells, brains, etc., not the reverse.

Whatever else it is or whatever impact it may have, the study of bacterial evolutionary relationships is central to the historical account of life on this planet. We may lay no claim to a comprehensive understanding of biology until we know this history, at least in its outline. And this is the perspective from which the present review is written.

### Three Ideas That Shape Our Concept of Bacterial Evolution

**Procaryote-eucaryote dichotomy.** The way we look at bacterial evolution is in essence shaped by a few picturesque, strongly held notions. One has to do with the place of bacteria in the spectrum of living systems; the other two concern our picture of how life began.

The prescientific distinction animal-vegetable-mineral is the starting point for our perception of the relationships among living things. Initially, every living thing was thought to be either a plant or an animal. The invention of the microscope, however, revealed a world of unicellular creatures, which because of their enormous and unusual variety caused us to wonder whether they were just very small animals or plants; there seemed to be another basic distinction, between macroscopic and microscopic forms, between multicellular and unicellular life. Haeckel's classic phylogeny, reproduced in Fig. 1, represents an amalgam of these two early views (76), in having three basic categories of living systems: plants, animals, and microorganisms (protists). This classical approach to phylogeny, based essentially on characteristics of the whole organism, has been refined in modern times to a five-kingdom scheme (139, 240, 241). However, such a taxonomy is not phylogenetically valid (see below).

It is intuitively evident that certain of an organism's characteristics, especially cellular attributes, are more essential than others. In the 1930s, E. Chatton (25) sought to construct a universal phylogeny on this principle by dividing the living world into two main groups, eucaryotes and procaryotes, on the basis of whether or not they possessed a true nucleus, i.e., one circumscribed by a nuclear membrane. (He also involved a few other intracellular structures found only in eucaryotes, such as the mitochondrion, to bolster the case [25].) Chatton's approach had clear virtues, and once its two basic categories were defined in detail, through electron microscopic studies and various molecular characterizations, the procaryote-eucaryote dichotomy became firmly (dogmatically) established as the primary phylogenetic distinction (138, 152, 209, 211).

However, the original definition of the procaryote carried an implication that no biologist recognized at the time, an implication that had profound consequences. Procaryotes were initially defined in a purely negative sense: they did not have this or that feature seen in eucaryotic cells. There was no logical reason to assume, therefore, that all procaryotes (all cells that were not eucaryotes, that is) were specifically related to one another. Yet, this is precisely what happened; "procaryote" was taken from the start to be a phylogenetically coherent taxon (209).

Over the years the definition of the procaryote (vis à vis the eucaryote) expanded from the initial negative one, based solely upon noncomparable characteristics, to a positive one, based upon comparable, molecular properties. In that process one procaryote, *Escherichia coli*, was assumed to represent all. Now we see the error in this never-tested assumption. Its unquestioned acceptance probably delayed the discovery of archaebacteria by well over a decade.

**Oparin Ocean scenario.** The old notion that the living world is somehow distinct from, unconnected to, the nonliving world (an idea that reflects creation myths, the mind-matter dichotomy, and other such things) gave rise to the panspermia notion. To counter both this idea and creationism, biologists were obliged to account for the origin of life in physical processes occurring on this planet. The first reasonably comprehensive attempt to do this was made by A. I. Oparin in the 1920s (159); a similar, but significantly different, proposal was made somewhat later by J. B. S. Haldane (77); and a slightly modified amalgam of the two is the only account of life's beginnings accepted by biologists today (104, 149, 160). The Oparin Ocean scenario, as it has been called, is frankly a "Just So Story." It is a vague and so not too useful hypothesis in its own right, and to the extent that biologists treat it as dogma, its scientific effectiveness is further diminished (249).

According to the current Oparin Ocean scenario, the primitive oceans became the ultimate repository for the great variety of chemicals and biochemicals thought to have been produced in the primitive anaerobic (at least, nonoxidizing) atmosphere through the action of ultraviolet light and electrical discharge upon water vapor, carbon dioxide, nitrogen, and other gases. In this way the primitive ocean became a "soup" of energy-rich biochemicals (104, 149, 160). Interactions among these produced ever more complicated structures, which eventually (somehow) turned into complex living (self-replicating) cellular entities (160). Since these earliest living systems arose in a bath of nutrients, they had no need to synthesize amino acids, nucleotides, and other products of intermediary metabolism and so did not develop such capacities. In other words, the first organisms were extreme heterotrophs, having neither photosynthetic nor autotrophic capabilities (160). They were in essence sinks for the chemical energy stored in the oceans (249). Only when these early cells began to exhaust their oceanic supply of nutrients did a need arise for intermediary metabolism, autotrophic capacity, and the ability to use light as an energy source, and only then did these features evolve, *deus ex machina* (88, 160).

This scenario has been interpreted by biologists to mean that the first organisms were (anaerobic) heterotrophic *procaryotes*, which later spawned photosynthetic and autotrophic sublines (and later still, procaryotes capable of aerobic metabolism). Textbooks customarily derive all living forms from an ancestral anaerobic fermentative heterotroph, taken to be some clostridium or streptococcus (15, 139). Eucaryotes are generally brought into this picture through a



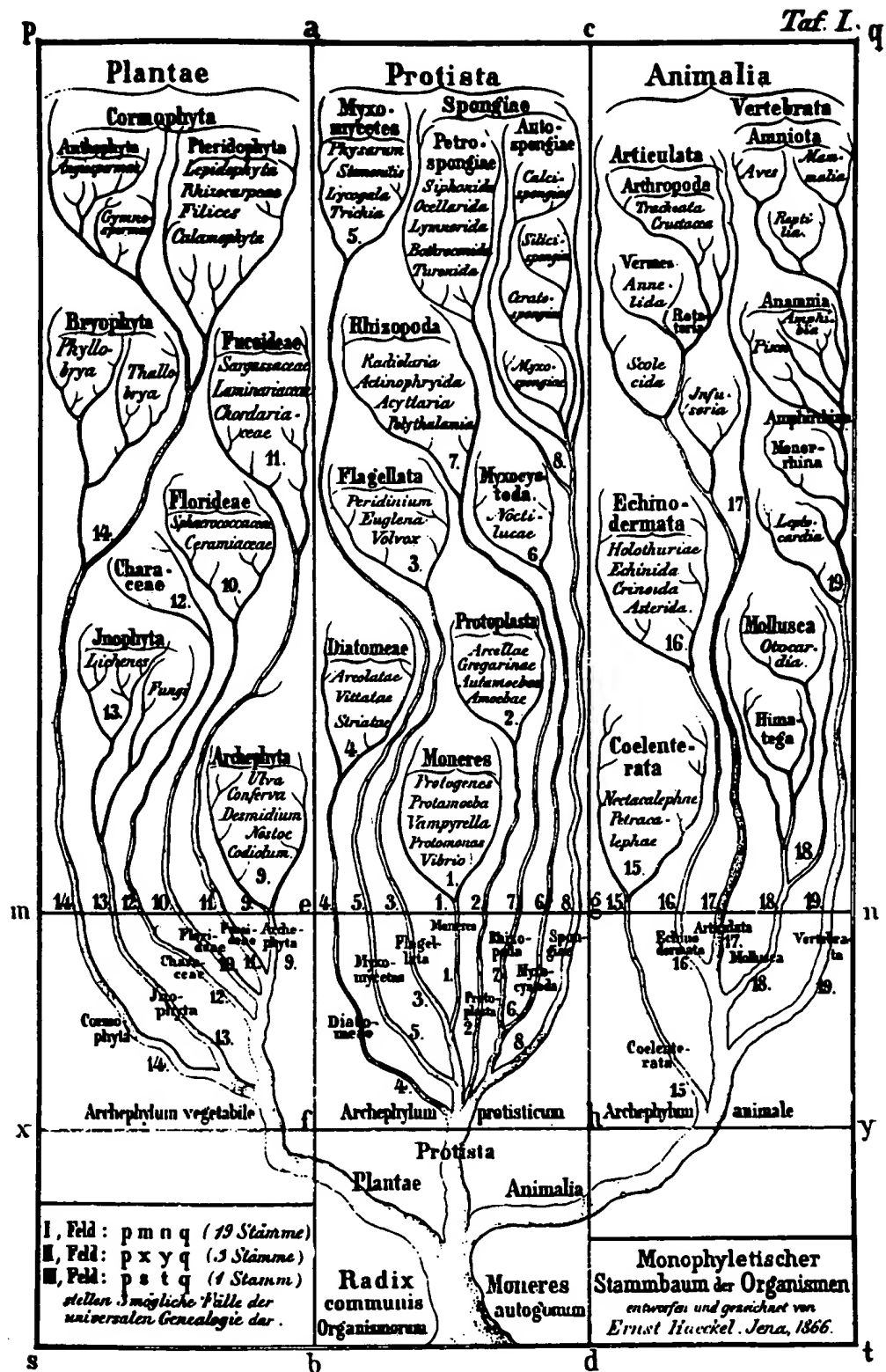


FIG. 1. Haeckel's phylogenetic tree of 1866 (76).

subline of wall-less (anaerobic) procaryotes that gained the capacity for endocytosis (138, 139). By ingesting other procaryotes that gave it additional metabolic capabilities (and molecular structures) and that eventually degenerated into organelles (contributing genes to the nucleus in the process), this primitive host cell, this urcaryote (256), became the eucaryotic cell.

Once again implicit assumptions have misled us, in this case the connotation of the prefix "pro-" in procaryote. Procaryotes had to precede eucaryotes; procaryotes are ipso facto older, simpler, and more primitive than eucaryotes and therefore gave rise to them. This is perhaps the main, if not the only, reason why the vast majority of biologists perceive the first organisms as procaryotic.

**Darwin's warm little pond.** A final, related element in the prejudices shaping our view of bacterial origins is Darwin's "warm little pond" image (31), which I am sure Darwin never intended to be a prescription for life's beginnings. (Darwin understood that the subject belonged to the future and probably intended more to dismiss it with this casual remark [31] than to give his successors a guiding principle.) Nevertheless, we are now stuck with this image of life's beginnings and have to cope with it as we do the other conceptual baggage we have inherited. Do microbiologists perhaps view thermophilic bacteria as adaptations from mesophilic species for this reason? "Warm" is an anthropocentrism. The setting in which bacteria arose may well have been warm, but it was not the hospitable warmth implicit in the pond Darwin pictured.

The collection of images associated with the procaryote-eucaryote dichotomy, the Oparin Ocean scenario, and, to a lesser extent, Darwin's warm little pond form the starting point and dictate the direction of our thinking about bacterial evolution. The basic flaws in all of them will become even more apparent as we proceed. It would be better, if that were possible, to forget about the lot and approach bacterial evolution with a clean slate. Molecular phylogenetic studies of bacteria are going to tell us a great deal, especially if our thinking is unfettered by old anthropocentric notions.

## MEASUREMENT OF BACTERIAL PHYLOGENETIC RELATIONSHIPS

### Three-Dimensional and One-Dimensional Characters

We intuitively recognize that complex characteristics (phenotypic patterns) are good indicators of relationship, i.e., and that a sufficiently complex pattern (character) is unlikely to have evolved more than once. However, our assessment that two or more organisms have the same or similar complex characteristics is by no means fail-safe. The judgment "similar" is too often subjective; what appears complex to our eye may not be so in the dynamics of the organism, and something that the biologist imagines as difficult to evolve may in reality be relatively simple (due to constraints on the system that he has not recognized, for example). Nowhere are our failings in this regard more evident than in the attempts to classify bacteria.

The sequencing of proteins and nucleic acids provides a new and more powerful approach to measuring evolutionary relationships and a new way of looking at them, in terms of the "evolutionary clock" (243). Genotypic information, i.e., sequence information, is superior in two main ways to phenotypic information, the classical basis for relating and classifying organisms: sequence information is (i) more readily, reliably, and precisely interpreted and (ii) innately

more informative of evolutionary relationships than phenotypic information is.

Unlike three-dimensional phenotypic patterns, a sequence pattern is one dimensional. One-dimensional patterns can be measured in simple ways, in terms of simple relationships. The elements of a sequence, nucleotides or amino acids, are restricted in number and well defined (quantized). The subjectivity that goes into the judgments "same," "similar," etc., at the phenotypic (three-dimensional) level is replaced by simple, more objective judgments and mathematically defined relationships in the world of sequences.

**The evolutionary clock.** The introduction of genetics into our model of the evolutionary process in the early part of this century was a major advance in that it let us understand evolution's "motor," the source of the variation upon which selection works. The concept of the evolutionary clock furthers this understanding; it shows us the relationship between this motor (i.e., genotypic change) and what we classically call evolution, the changes in phenotype. At the level of the genotype, change constantly occurs. However, most of it is of a nature that it is not acted upon by selection (105, 106). It therefore, becomes fixed randomly in time, making its characterization as "clocklike" in occurrence appropriate. In other words, evolution has a tempo that is quasi-independent of its mode (the selected changes occurring in the phenotype). An analogy to a car and its motor is apt: a car does not go unless its motor is running, but the motor can run without the car moving.

Cytochrome *c* evolution provides a good example of the evolutionary clock. An enormous number of different versions of this sequence all appear to be equivalent functionally (44). (Formally speaking, the mapping from genotype to phenotype [upon which selection acts] is degenerate.) A change from one such version to another would then occur randomly, independent of selection; the probability of its occurrence would only reflect a lineage's mutation rate (105, 106). Since the number of possible functional configurations for a given gene is enormous by any standards, similarity at the genotypic level (i.e., extensive sequence homology) can never reflect convergent evolution. Consequently, cytochrome *c* sequence comparisons have been used very successfully to time key events in eucaryotic evolution (in sequence distance terms) and to determine molecular genealogies. Phylogenetic trees based upon cytochrome *c*, or similar molecular chronometers, represent significant improvements over their classical counterparts based upon phenotypic comparisons (49).

**A missed opportunity.** The now classic paper of Zuckerkandl and Pauling, in 1965, effectively launched biology into the world of molecular chronometers (283). Although they may not have been the first to recognize that sequence comparisons could be used to define phylogenetic relationships (protein-sequencing technology had been in existence for about a decade by then), Zuckerkandl and Pauling were the first to put the case in a well-defined, scientifically effective way. This was precisely what the microbiologist needed to resolve the problems of the natural relationships among bacteria. However, microbiologists did not rush to utilize the new approach; remember, they had by then come to see the problem as unimportant. Nevertheless, a peripheral sort of awareness existed; the new molecular techniques were perceived by some as useful in standard bacterial classification, and a number of small-scale efforts were mounted to improve classification by various molecular approaches. Genetic characterization, deoxyribonucleic acid (DNA) base ratios, nucleic acid hybridization studies,

cell wall analyses, and (a little) protein sequencing began to reveal phylogenetically valid groupings (4, 140, 183, 188).

These early molecular approaches, though useful, were not powerful enough to reveal the higher bacterial taxa, and in any case conventional wisdom did not perceive doing so as important. While gram-positive cell walls exhibited interesting and informative variety in their composition, the gram-negative ones were too uniform to be of much use in defining taxa (183). Nucleic acid hybridization work, i.e., DNA/DNA studies, were (necessarily) confined to relationships within genera (185, 207). When DNA/ribosomal ribonucleic acid (rRNA) hybridization studies were instituted, they were used only to revise existing local taxonomic structure (32, 33). Even protein sequence comparisons provided no insights, except in one particular group of purple bacteria (4, 148).

#### Nature of Molecular Chronometers

A molecule whose sequence changes randomly in time can be considered a chronometer. The amount of sequence change it accumulates (formally a distance) is the product of a rate (at which mutations become fixed)  $\times$  a time (over which the changes have occurred). The biologist cannot measure this change, however, by comparison of some original to some final state, since the original state (ancestral pattern) is not accessible to him. Instead, he uses the fact that two (or more) versions of a given sequence that occur in extant representatives of two (or more) lineages have ultimately come from the same common ancestral pattern, and so measures the sequence difference between the two (or more) extant versions, which is roughly twice the amount of change that each lineage has undergone (assuming comparable rates of change in each) since they last shared a common ancestor.

All sequences are not of equal value in determining phylogenetic relationships. To be a useful chronometer, a molecule has to meet certain specifications as to (i) clocklike behavior (changes in its sequence have to occur as randomly as possible), (ii) range (rates of change have to be commensurate with the spectrum of evolutionary distances being measured), and (iii) size (the molecule has to be large enough to provide an adequate amount of information and to be a "smooth-running" chronometer (explained below)).

**Clocklike behavior.** A molecular chronometer should measure, should be representative of, the overall rate of evolutionary change in a line of descent. One might think that the best chronometer would then be a genetic segment upon which there are no selective constraints. Its changes would occur randomly along the length of the segment, occur in a quasi-clocklike fashion, become fixed at a rate equal to the lineage's mutation rate (105, 106), and be easy to interpret. However, such sequences are of little value for phylogenetic measurement, because they generally do not meet the second requirement; their rates are so rapid that they cover only very restricted phylogenetic ranges (38). Such sequences are evolutionary stopwatches; they measure only the short-term evolutionary events. The third (i.e., degenerate) codon positions in structural genes are often used in this capacity (20, 128, 243).

The more useful molecules for phylogenetic measurement all represent highly constrained functions. Some sequences of this type change slowly enough to span the full evolutionary spectrum. Unfortunately, what makes them useful as chronometers also makes them problematical. Strict clocklike behavior is usually hard to find, i.e., identify. Unless

functional constraints remain strictly constant over the evolutionary range being covered, nonrandom (selected) sequence changes accumulate, over and above the randomly introduced ones, and artificially increase phylogenetic distances between organisms, which usually leads to improper determinations of branching orders. Cytochrome *c* provides an example. In the  $\alpha$  subdivision of the purple bacteria (defined below) the molecule changes in size, from "medium" to "large" (37), reflecting some unknown and subtle functional change. This size change has brought with it additional nonrandom sequence changes that appear to distort the phylogenetic determination somewhat (4). A similar situation can be seen in the phylogeny of certain *Bacillus* species constructed from 5S rRNA sequence comparisons (87).

A second problem with highly constrained chronometers is the extremely different rates at which the various positions in a sequence tend to change. This in itself is not a problem; the hands of a clock move at very different rates. However, analysis of the data becomes difficult at this point (see below).

**Phylogenetic range.** The world of bacterial evolution is vast in comparison to that of the eucaryotes with which we are familiar, i.e., the metazoa. A billion years is a relatively short time in bacterial evolution. Therefore, the range of chronometers used to measure phylogenetic relationships among bacteria needs to be considerably greater than what is optimal for metazoa. Cytochrome *c* is an excellent chronometer for measuring much of eucaryote phylogeny (49). Among the eubacteria, its effective range is restricted to the subdivision level; it orders the  $\alpha$ -purple bacteria (see below), but does not relate these accurately to any other subdivision of the purple bacteria (4, 148).

**Size and accuracy.** In addition to the obvious need for large size in a chronometer (i.e., good statistics), there is a more subtle one. Size per se is probably not what is important. It is that the molecule consists of a fairly large number of loosely coupled "domains" (functional units), regions that are somewhat independent of one another in an evolutionary sense (250). In this case, nonrandom changes affecting one of the units will not appreciably affect the others; therefore, when one part of the chronometer becomes drastically altered by introduction of selected changes (i.e., gives a distorted reading), the other parts remain practically unaffected. The more units of this kind a chronometer contains, the less sensitive are its measurements of evolutionary distances to nonrandom changes in one of them, i.e., the "smoother" the chronometer runs. This is a major difference between 5S rRNA, for example, and the large rRNAs (250).

#### rRNAs, the Ultimate Molecular Chronometers

Why they are so good. rRNAs are at present the most useful and most used of the molecular chronometers. They show a high degree of functional constancy, which assures relatively good clocklike behavior (250). They occur in all organisms, and different positions in their sequences change at very different rates, allowing most phylogenetic relationships (including the most distant) to be measured, which makes their range all-encompassing. Their sizes are large and they consist of many domains. There are about 50 helical stalks in the 16S rRNA secondary structure and roughly twice that number in the 23S rRNA (75, 155), which makes them accurate chronometers on two counts.

Perhaps the most compelling reason for using rRNAs as chronometers is that they can be sequenced directly (and,

therefore, rapidly) by means of the enzyme reverse transcriptase (116, 174). This distinguishes them from all other potential chronometers in the cell except for a few of the smaller RNA species (which do not have as good chronometric properties, and in some cases cannot be isolated with the ease with which rRNAs can). It is reasonable for a properly equipped laboratory in the future to sequence on the order of 100 16S rRNAs per year.

**Oligonucleotide cataloging.** Until several years ago it was not feasible to determine complete rRNA sequences. So more than a decade was spent characterizing them in terms of partial sequences, by the so-called oligonucleotide cataloging method. Short oligonucleotides, of lengths up to 20 or so bases, are produced by digestion of 16S rRNAs with ribonuclease T<sub>1</sub> (which cleaves specifically at G residues) (55); a collection of these sequence fragments from a given rRNA constitutes an oligonucleotide catalog: a detailed, complex pattern characteristic of a given bacterial species (55). Comparisons among these catalogs permits phylogenetic groupings to be identified at various taxonomic levels, including the highest (55, 257).

Oligonucleotide catalog data are usually analyzed in terms of binary association coefficients, so-called  $S_{AB}$  values (55), defined as the ratio of twice the sum of bases in oligonucleotides (length greater than five) common to two catalogs *A* and *B*, to the sum of all bases (in oligonucleotides of length greater than five) in the two catalogs (55). The relationship between  $S_{AB}$  value and percentage sequence similarity cannot be theoretically derived (because the relative rates at which individual positions change are not predictable *a priori*). A plot of percent similarities and corresponding  $S_{AB}$  values for a collection of 16S rRNA sequences is shown in Fig. 2.  $S_{AB}$  is seen to vary approximately as the fifth to sixth power of percent similarity. The relationship between the two measures is clearly not a precise one, especially below  $S_{AB}$  values of about 0.40.

While the cataloging method sufficed to define most of the major bacterial phyla, it generally failed to resolve the branching orders among them or among their subdivisions. (Such distinctions are never easily made, as evidenced by the fact that the animal phyla have been known for a century, but the order of their branching from one another has yet to be determined.) The cataloging approach also ran into difficulties over the branching order of rapidly evolving lines of descent, again a perennial problem. Full sequencing of 16S rRNA has now replaced the earlier oligonucleotide cataloging approach, a development that greatly increases the resolving power of the rRNA chronometer.

#### Analysis of Sequence Alignments

Given a sequence alignment, which for rRNAs can be constructed in a straightforward empirical manner (75, 260), the question becomes how to analyze it, i.e., how to extract the most phylogenetically useful information. At present three main methods are used for this purpose: (i) distance matrix treeing, (ii) maximum parsimony analysis, and (iii) cluster analysis.

**Distance matrix methods.** Distance matrix methods (47, 49, 157) utilize only the sequence distances between pairs of sequences, i.e., the fraction of positions in which the two sequences differ. This distance is actually an underestimate of the true evolutionary distance between sequences; although most of the differences between two sequences reflect single mutational events at any given position in a

sequence alignment (when the chronometer is working in its effective range, that is), some of them represent multiple events. Were all lineages and all positions in a sequence changing at the same rate, then correction for this effect, conversion of sequence distances to evolutionary distances, would be a relatively simple matter (99). Unfortunately, this is not the case, either for different lineages or for different positions in the sequence, and the proper correction for multiple changes remains a major problem in tree construction.

Distance matrix treeing assumes that evolutionary distances conform to a tree topology. To use a simple example, let AB, AC, AD, BC, BD, and CD be the six determined evolutionary distances among four sequences, A, B, C, and D. If the three pairwise sums AB + CD, AC + BD, and AD + BC meet the condition that two of them are equal and greater than or equal to the third, a condition that can be understood by reference to Fig. 3, then the data fit a tree topology. When this condition is met even approximately, evolutionary distances can be used to reconstruct phylogenies with reasonable accuracy (36, 47, 49, 157).

Given a matrix of (corrected) evolutionary distances for a sequence alignment, one in principle examines all possible phylogenetic trees (branching orders), treating branch lengths as adjustable parameters, and declares the one that best fits the data (by a least-squares analysis) to be the "correct" tree (47, 49, 157). This, of course, is not how things are actually done, for the number of all possible branching orders rapidly becomes computationally unmanageable as the number of sequences in the alignment reaches even a moderate number. There are many approximations, many competing algorithms for giving the "best" tree in a reasonable amount of time (47, 49, 157).

**Maximum parsimony analysis.** Unlike sequence distance matrix analysis, maximum parsimony analysis does not reduce the differences among sequences to a single number, a distance; it treats the positions individually (47, 48). Its ~~assumption is that the correct phylogenetic tree is the most parsimonious one, the one for which the smallest overall number of mutational changes have to be postulated to arrange the set of considered sequences upon it.~~ (Each branch in the tree, each segment between branching points, is defined by the specific changes that occur [in some ancestral sequence] on that branch.) As with distance matrix treeing, one in principle looks at all possible tree branching arrangements and chooses the most parsimonious one (47, 48). Also, as with distance matrix treeing, the problem of finding the correct tree can be computationally intense, even more so in this case, and much time and effort have gone into devising computer algorithms that do this as efficiently as possible (46-48).

**Cluster analysis.** Cluster analysis, the third major method of analyzing sequence data, ~~groups sequences on the basis of how similar they are to one another or to other groups of sequences (193, 194). The method is the least computationally intense of the three, but also the least accurate.~~

The main difficulty in all analyses of sequence data lies in the fact that different lineages and different positions in a sequence can evolve at significantly different rates. Making distance corrections on the assumption that all positions in a sequence change at the same rate (65, 99, 157) underestimates the correction needed. Parsimony analysis tends not to position rapidly evolving lineages correctly and is confused by rapidly changing positions, perhaps more so than is distance treeing (157). Cluster analysis is especially sensitive to these problems; rapidly evolving lineages are as a rule

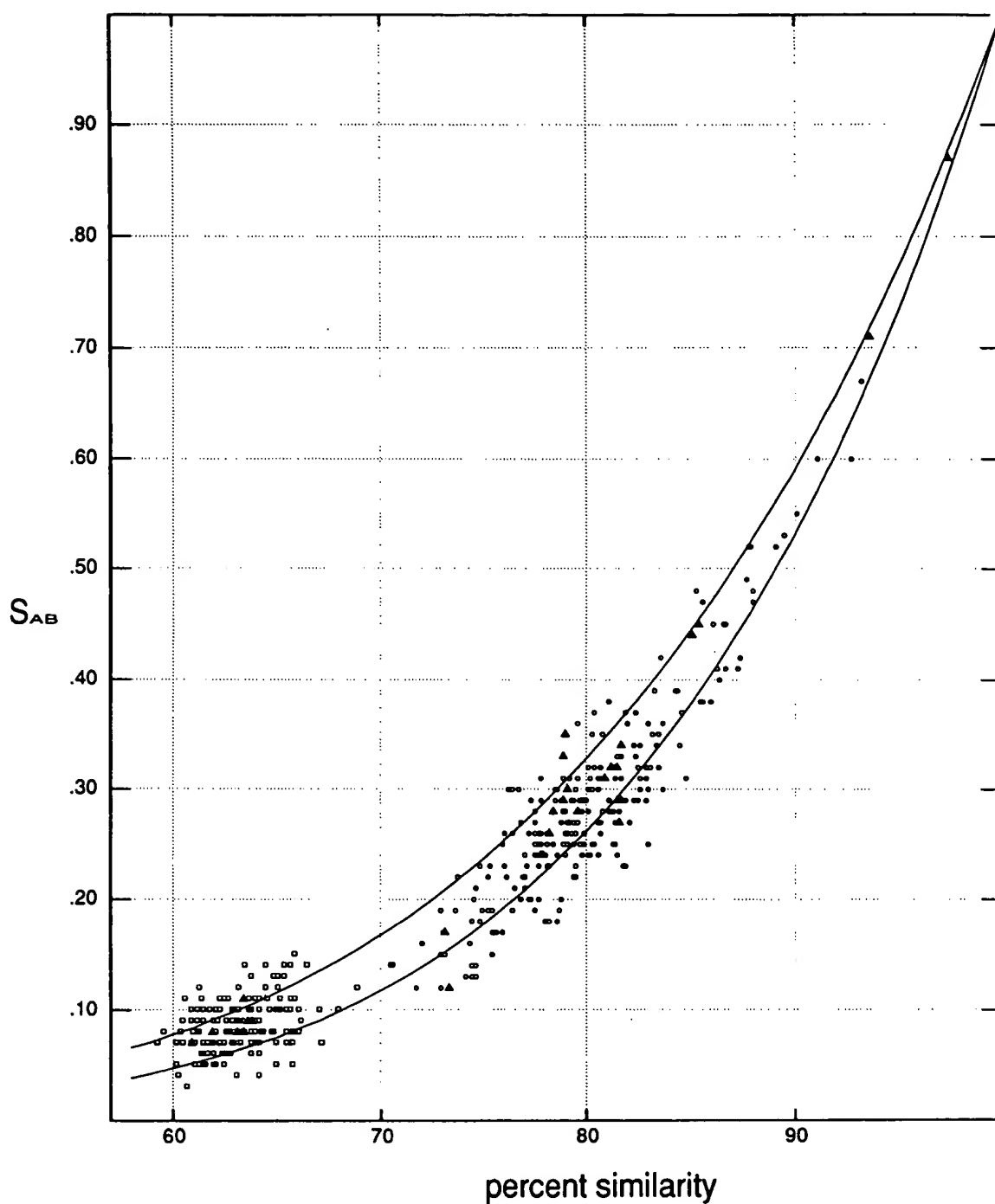


FIG. 2. Plot of percent sequence similarity versus binary association coefficient ( $S_{AB}$  value [55]), for a representative sampling of eubacterial and archaeobacterial sequences (unpublished analysis). The two theoretical curves are  $X^2$  (upper curve) and  $X^6$  (lower curve), where  $X$  = percent similarity. Symbols: O, values for pairs of eubacterial sequences; □, values for eubacteria with archaeobacteria; ▲, values for *E. coli* with either eubacteria or archaeobacteria.

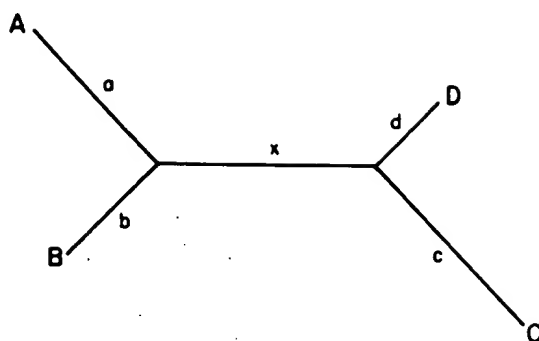


FIG. 3. Unrooted tree for four species, A, B, C, and D, illustrating the relationship that must hold among the six evolutionary distances,  $AB = a + b$ ,  $AC = a + x + c$ ,  $AD = a + x + d$ ,  $BC = b + x + c$ ,  $BD = b + x + d$ , and  $CD = c + d$ , for them to fit a tree topology (see text).

positioned too deeply in trees so constructed. In other words, analyses of sequence data today are far from optimal.

The needed improvements in analysis will not come primarily from better theoretical treatments or more efficient algorithms per se. Improvement will be basically the result of empirical approaches. Given a large enough sequence data base, it will become possible to describe the pattern (rate and kind) of change at given positions in a (rRNA) molecule and design specific analyses based upon this description, analyses that correct more accurately for multiple changes at sites and utilize only those positions in the molecule appropriate to the phylogenetic range being measured. (The optimal method will not consider the second hand when timing the seasons.)

Consideration of the detailed chronometric structure of rRNA will be postponed until the reader has some familiarity with bacterial phylogeny. However, the reader who already has this familiarity may wish to read that discussion (which begins on page 253) at this point.

#### DO BACTERIA HAVE A GENEALOGY AND A MEANINGFUL TAXONOMY?

In classifying bacteria microbiologists make two implicit assumptions: (i) that bacteria have a phylogeny, and (ii) that the taxonomic system that works well for the metazoa is actually applicable to, i.e., meaningful in, the microbial world. These two points require explication and discussion, for they are far from self-evident.

In questioning the first assumption, one at the very least questions approaches to measuring bacterial genealogies and beyond that whether in principle the bacterium as a whole has a genealogy, a unique history. These questions are raised by the existence of lateral (interspecies) transfer of genetic information among bacteria (23, 177). A given gene (or set of genes), say for nitrogen fixation in *Azotobacter* sp., might have been evolved in an organism not immediately related to *Azotobacter* and have been acquired by that organism through plasmid transfer or some similar process. When used as molecular chronometers such genes would not yield the correct genealogy for *Azotobacter*. In the extreme, interspecies exchanges of genes could be so rampant, so widespread, that a bacterium would not actually have a history in its own right; it would be an evolutionary chimera, a collection of genes (or gene clusters), each with its own history.

Fortunately the matter is experimentally decidable. Were an organism an evolutionary chimera, then its various chronometers would yield different, conflicting phylogenies. A limited test of the possibility can be made for the  $\alpha$  subdivision of the purple bacteria, for which a number of species have been characterized by both rRNA catalogs and cytochrome *c* sequences. Phylogenetic trees derived from the two molecules have nearly the same topology, strongly suggesting that neither chronometer has been involved in interspecies gene transfer (258). Although more extensive testing of the lateral transfer notion is highly desirable, it is now relatively safe to assume that bacteria do in principle have unique, characteristic evolutionary histories and that at least some of the cell's chronometers record them. (What is not known is the fraction of the functions in a bacterial cell that are subject to interspecies gene transfer, and which ones these are.)

Given that a bacterial genealogy exists, the question (the second assumption above) then becomes whether such a cladogram can be divided into zones (into taxa) that are naturally, as opposed to artificially, defined and, if so, whether the groupings arrange themselves naturally into simple hierarchical structures (into taxonomic levels).

The metazoan (Linnaean) classification system, though still imperfectly implemented, is as useful as it is because metazoa (chiefly animals) intrinsically group into (naturally defined) categories. The metazoan kingdoms and the animal phyla (192), for example, are readily differentiable groupings. And, although fine points are debated, a metazoan species is also well defined, largely because of the constraints imposed by mating (144). A bacterial species is certainly far more problematical a concept than a metazoan species (8, 29). In the present context, however, our concern is essentially with the higher taxa: whether or not (within a bacterial urkingdom) these are somehow naturally defined or are mere artificial constructs.

There is no compelling evidence to suggest that the bacteria fall into naturally defined taxa. In fact, existing evidence might even suggest the contrary. Few of the (extensively investigated) bacterial phyla presented below can be defined by phenotypic properties common to all members of the group. For example, the gram-positive bacteria defined by cell wall structure form a clade, but this clade also includes bacteria that do not have gram-positive walls. Although the purple bacterial group is named for the particular type of photosynthesis done by some of its members, the photosynthetic pigment does not define the group as a whole, which also includes many nonphotosynthetic species.

Nevertheless, I feel that ultimately bacteria will be shown to fall into naturally defined taxa. One reason this is not obvious at present may be that various bacterial groups have been studied from different perspectives: what we know to be characteristic of one may never have been looked for in another. This fact alone could explain some of the apparent lack of phenotypic resemblance among genealogically clustered species. Another reason is that the microbiologist has never before had phylogenetically defined groupings that he could count on to direct his search for phenotypically unifying characters. Recent studies utilizing such an approach (97) appear promising.

The main reason for thinking that bacterial taxa are naturally defined is that the characteristics of the rRNA chronometer (discussed below) strongly suggest this. Under certain circumstances rRNAs will accumulate unusual sequence changes, ones that normally occur with a negligible

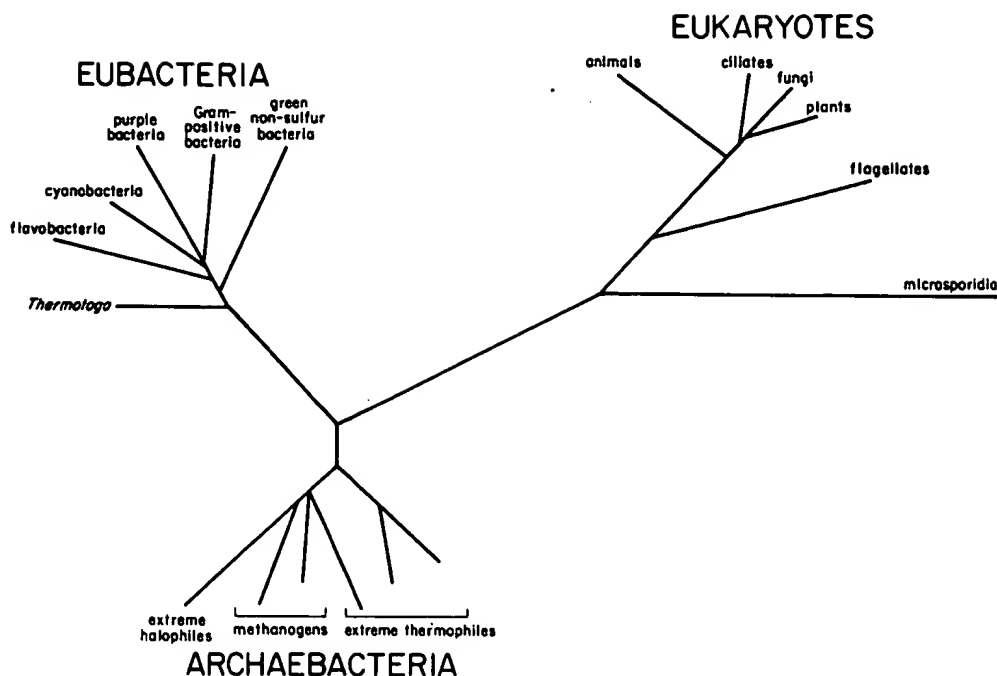


FIG. 4. Universal phylogenetic tree determined from rRNA sequence comparisons. A matrix of evolutionary distances (99) was calculated from an alignment (260) of representative 16S rRNA sequences from each of the three urkingdoms. This was used to construct a distance tree (36), based upon those positions represented in all sequences in the alignment in homologous secondary structural elements (75). Line lengths on the tree are proportional to calculated distances. The alignment includes the following eubacterial sequences: *Thermotoga maritima* (1); green non-sulfur bacteria, *Thermomicrobium roseum* (162); flavobacteria, *Flavobacterium heparinum* (234); cyanobacteria, *Anacystis nidulans* (224); gram-positive bacteria, *Bacillus subtilis* (68); and purple bacteria, *Escherichia coli* (19); the following archaeobacterial sequences: extreme halophiles, *Halobacterium volcanii* (72); methanogens, *Methanococcus vannielii* (96) and *Methanobacterium formicicum* (124); and extreme thermophiles, *Thermococcus celer* (Woese et al., unpublished data), *Desulfurococcus mobilis* (R. Garrett, personal communication), and *Thermoproteus tenax* (126); and the following eucaryotic sequences: microsporidia, *Vairimorpha necatrix* (226a); flagellates, *Euglena gracilis* (196); cellular slime molds, *Dictyostelium discoideum* (145); ciliates, *Paramecium tetraurelia* (195); fungi, *Saccharomyces cerevisiae* (179); plants, *Zea mays* (147); and animals, *Xenopus laevis* (181). Branching order within each kingdom is correct to a first approximation only. See the trees for the individual kingdoms for precise branching orders.

frequency. These unusual changes seem to accompany the formation of various major branches on the tree, which correspond to major shifts in bacterial phenotype. The reason for this lies in the tempo-mode relationship. Thus the rRNA chronometer may provide a means of defining natural bacterial groupings that is purely genotypic, independent of any phenotypic definition thereof.

#### THE UNIVERSAL PHYLOGENETIC TREE

Although we evolutionists still have much to learn a century after Darwin's death, a major milestone has recently been reached. Molecular phylogenetic approaches let us for the first time see the full extent of the tree that encompasses all extant life (72, 263); see Fig. 4. As yet this tree must be drawn in an unrooted form, because the crucial question of the point in its structure that corresponds to the Universal Ancestor, the point from which all extant life ultimately emanates, remains unanswered.

Perhaps the most striking characteristic of the universal tree is the distinctness of the primary kingdoms, the large sequence distances that separate one kingdom from another. Figure 5, a plot of percent sequence similarities for an

extensive collection of procaryotic 16S rRNAs, demonstrates this in graphic form; every eubacterial sequence is far closer to every other eubacterial sequence than to any archaeobacterial sequence, and vice versa. (The interkingdom distances in Fig. 4 are even best considered lower bounds, for they are large enough that they could well have been underestimated.)

The extent of sequence distance that separates the primary kingdoms is reflected in the degree of phenotypic difference among them. It has long been obvious that eucaryotes are quite distinct from procaryotes (i.e., eubacteria); the two differ in general cellular organization, in genome structure, in control and expression of genetic information, in the structure of the translation apparatus, and in the details of numerous molecular functions. In the last decade we have found the same to be true for the archaeobacteria. They too have unique molecular architecture, cellular organization, genome structure, etc. (see assorted chapters in references 102 and 270). The degree of genotypic and phenotypic separation among the primary kingdoms argues that the ancestor they all shared was a special sort of entity (251), whose nature will be discussed below.

Each of the primary kingdoms has its particular form of

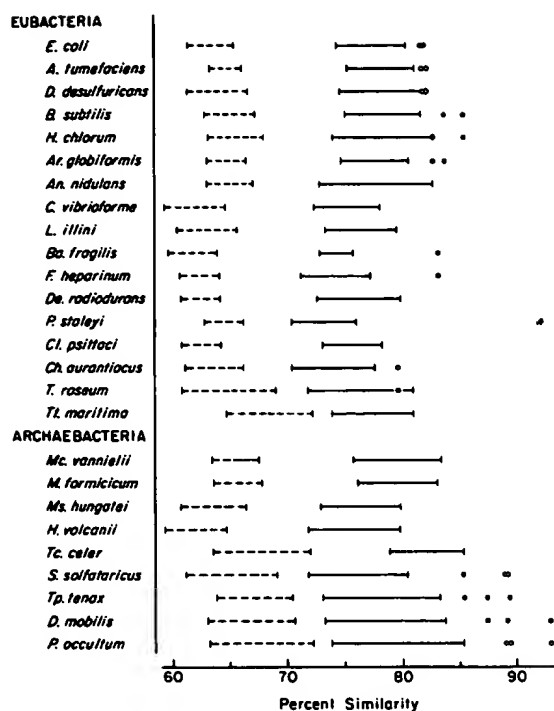


FIG. 5. Percent sequence similarity for various eubacterial and archaeobacterial 16S rRNAs. Open circles indicate percent similarity for a given sequence with others in the same phylum. The range of similarities between a given sequence and those from other phyla in the same kingdom is shown as a solid line, whereas the corresponding range across kingdom lines appears as a dotted line. Only those positions represented in all sequences in the composite alignment (in homologous structures) were used in the calculations.

rRNA. Secondary structures for the three types of 16S-like rRNAs can be seen in Fig. 6. While there are general resemblances among them, there are also characteristic differences in structural detail (75).

Strong rRNA sequence signatures, i.e., positions in the molecule that have a highly conserved or invariant composition in one kingdom, but a different (highly conserved) composition in one or both of the others, also define and distinguish the three urkingdoms. Table 1 is a signature distinguishing archaeobacteria from eubacteria; the location of these positions in relation to the molecules' secondary structure can be seen in Fig. 7. Though not presented in this review, equally pronounced signatures exist for eucaryotes (75, 260).

Archaeobacteria can resemble either eubacteria or eucaryotes (or neither), depending upon what phenotypic characters are considered. For example, the 16S rRNAs of the two procaryotic kingdoms are relatively close in structure (75), whereas 7S RNA sequence and structure reveals a resemblance between archaeobacteria and eucaryotes (129, 134, 151; B. P. Kaine, unpublished results). Eucaryotes and eubacteria have similarities as well, e.g., ester-linked lipids (117, 118), but these are relatively few. Our present, rather limited understanding would suggest that the overall phenotypic resemblance is greatest between the archaeobacteria and eucaryotes.

What the biologist, especially the microbiologist, must now fully recognize is that there no longer exists any reason to consider that archaeobacteria and eubacteria are related to one another to the exclusion of eucaryotes. Unfortunately, the title of this review, "Bacterial Evolution," implies the opposite. The title is intended, however, merely as a celebration of the fact that within the last 10 years the field of bacterial evolution has passed from a suspect discipline, about which almost nothing was known, to a full-fledged area of scientific investigation, rich in its implications for all of biology.

## EUBACTERIAL PHYLOGENY

### Background

Not only did we know very little about eubacterial phylogeny before the advent of the rRNA approach, but what we thought we knew tended to be wrong. The old idea (justified to some extent by cell wall compositions [183]) that there are two primary categories of (eu)bacteria, gram positive and gram negative, turns out to be a half-truth. Gram positive is indeed a phylogenetically coherent grouping, but gram negative is not. The latter encompasses of the order of 10 distinct groups, each the equivalent of the gram-positive one, as we will see. The old idea that wall-less bacteria, mycoplasmas, are phylogenetically remote from other (eu)bacteria (62, 176) is incorrect; the true mycoplasmas are merely "degenerate" clostridia (see below). Photosynthetic bacteria do not form a grouping genealogically distinct from the nonphotosynthetic bacteria (92, 168, 169); actually, the major photosynthetic types each represent separate high-level phylogenetic units which in most cases include many nonphotosynthetic species as well (56, 64, 206, 266). Autotrophs and heterotrophs are not phylogenetically separate groupings; they are intimately intermixed within the various eubacterial phyla (56, 206). Thus, the textbook views in which photosynthetic (or autotrophic) bacteria arise from nonphotosynthetic heterotrophic ancestors (15, 139) gain no support from the rRNA-based phylogeny. It is the classical microbiologist's insistence on morphology as the primary criterion (108), a prejudice inherited from the botanist, that more than anything engendered the confused and confusing state of bacterial taxonomy; almost none of the taxa (beyond the level of genus) defined primarily in this way pass phylogenetic muster (56).

The misconceptions of the classical microbiologist cannot be condemned. They were innocent attempts to create a phylogenetic framework at a time when phylogenetic classification was not experimentally possible. Unfortunately, many of them are now formalized as accepted bacterial taxonomy: they are repeated over and over in textbooks. They shape the design and interpretation of our experiments. And, because this taxonomy is patterned after metazoan classification, it is a de facto phylogenetic statement. Such an entrenched, organized system will not change easily. Yet replacing the old taxonomy with a phylogenetically valid one is vital to the future development of microbiology.

### Major Eubacterial Groups and Their Subdivisions

As of this writing over 500 species of eubacteria have been characterized in rRNA terms. Although most of these characterizations utilized the older and less informative oligonucleotide cataloging method, the more than 50 nearly complete sequences now known are sufficient in number that the



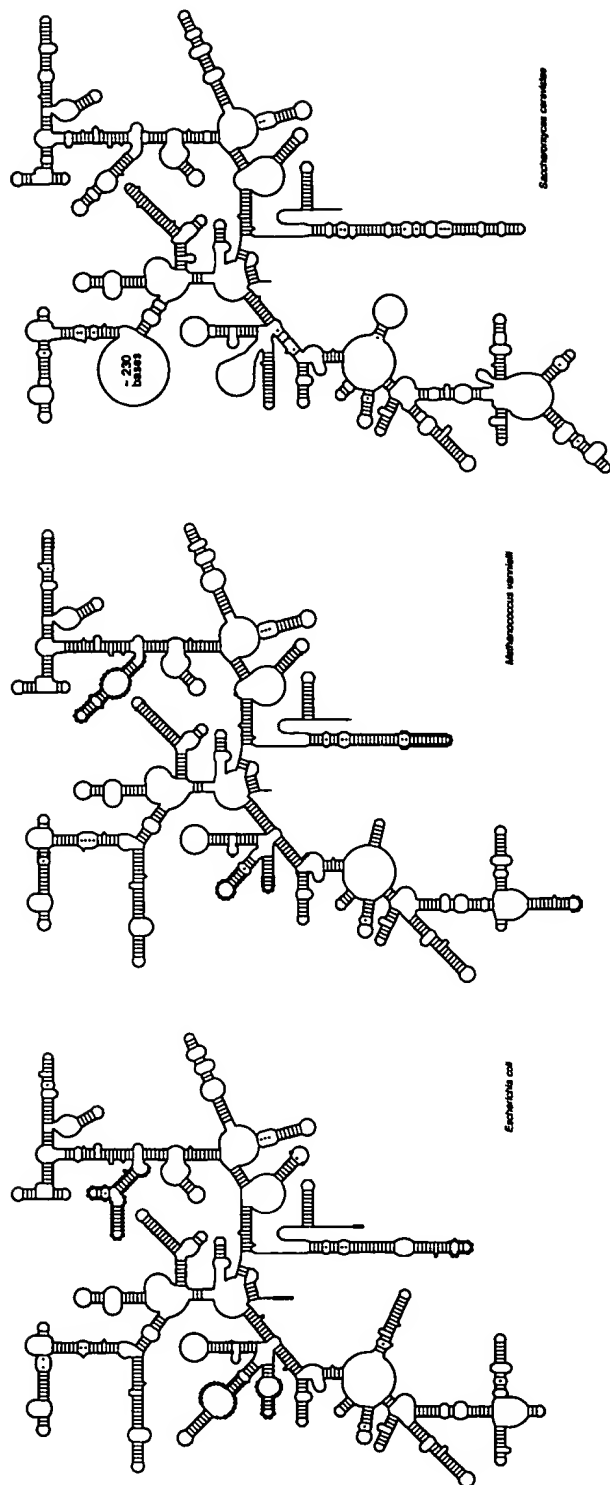


FIG. 6. Representative small-subunit rRNA secondary structures for the three primary kingdoms (75). Dots in the *Escherichia coli* or *Methanococcus vannielii* structures identify the positions or locales where eubacteria and archaeobacteria characteristically differ.

TABLE 1. Sequence signature distinguishing the two procaryotic kingdoms<sup>a</sup>

Position <sup>b</sup>	Composition			Position	Composition		
	Eubacteria	Archaeobacteria	Eucaryotes <sup>c</sup>		Eubacteria	Archaeobacteria	Eucaryotes
9	G <sup>d</sup>	C	C	585	G	C	U
10	A	Y	U	675	A	U	U
24	U	R	A	677	u	C	Y
25	C	G	G	684	u	G	G
31	g <sup>e</sup>	—	—	707	U	C	Y or G
33	A	Y	A	716	A	C	Y
38	G	A	u	756	C	G	A
39	G	U	g	862	Y	G	a
43	C	R or U	Y	867	G	C	u
47	y	R	R	912	C	U	U
53	R	C	C	923	A	G	A
113	G	C	C	930	C	A	G
314	c	G	G	931	C	G	G
338	A	G	A	934:1	—	Y or A	c
339	C	G	C	952	U	C	C
340	U	C	N	962	C	G	U
349	A	G	R	966	G	U	U
350	G	Y	G	973	G	C	G
358	Y	G	G	975	A	G	G
361	R	C	C	1045	c	G	R
367	U	C	U	1060	U	C	C
377	G	C	Y	1086	U	C	Y
386	C	G	R	1087	G	C	U
393	A	G	A	1109	C	A	A
397	A	—	—	1110	A	G	G
399	G	Y or A	R	1194	U	G	R
403	C	A	A	1197	A	G	G
508	U	C	A	1211	U	G	Y
514	Y	G	G	1212	U	A	a
523	A	C	A	1229	A	G	G
537	R	C	C	1381	U	C	C
540	g	C	C	1386	G	C	C
549	C	U	c	1387	G	U	C
551	U	R	U	1393	U	C	U
558	G	Y or A	A	1415	g	C	C
559	A	U or G	N	1485	u	G	g

<sup>a</sup> Based upon approximately 30 eubacterial and 12 archaeobacterial 16S rRNA sequences (19, 68, 224, 263, 274; C. R. Woese et al., unpublished data). Y, Pyrimidine; R, purine; N, any base; —, no base.

<sup>b</sup> Position in 16S rRNA sequence.

<sup>c</sup> Eucaryotic compositions taken from alignment used to construct Fig. 14; upper case, no exceptions; lower case, one exception.

<sup>d</sup> Upper case, Base found in all or all but one sequence in the kingdom but in no more than one sequence in the other kingdom.

<sup>e</sup> Lower case, As for footnote d but with two to three exceptions.

conclusions derived from the older approach can be significantly refined and extended.

At the level of oligonucleotide analysis it was apparent that the bacteria separate into more or less naturally defined "phyla" (see discussion below). This was not apparent from the binary association coefficients ( $S_{AB}$  values), but could be seen in oligonucleotide signatures (266). (Such signatures are collection of specific oligonucleotides, each of which occurs in most or all members of a given phylogenetic group but rarely, if at all, in most other groups, especially closely related ones [266].) Full rRNA sequences now permit the identification of individual positions in the molecule, sequence signatures that define the various groupings. We will use these rather than the older oligonucleotide signatures in the discussion below.

In some cases taxa defined by rRNA sequences can be identified phenotypically by common characteristics of the group. For the majority of such taxa, however, any given common character will link most, but not all, members of the

group. A few of the rRNA groups are without phenotypic justification. The spirochetes and relatives (see below) seem to be a taxon of the first type; all identified members of the genotypically defined unit possess the unique and unusual axial fibrils classically associated with these organisms (24). The gram-positive phylum is an example of the second type; the majority of its species have characteristic gram-positive cell walls (183), but a few lineages, such as *Helicobacterium* (255) and the mycoplasmas (176), do not have them. An example of a genotypically defined unit for which no convincing phenotypic justification can be given is the bacteroides-flavobacter phylum (166, 234); see below. In this case and others like it, the lack of common phenotypic properties could merely reflect the fact that the various subgroups have been characterized in entirely different ways and so would not necessarily be expected to show common characteristics. Such groups challenge the microbiologist to discover their unifying phenotypic motifs.

The eubacterial phyla and their subdivisions as they are

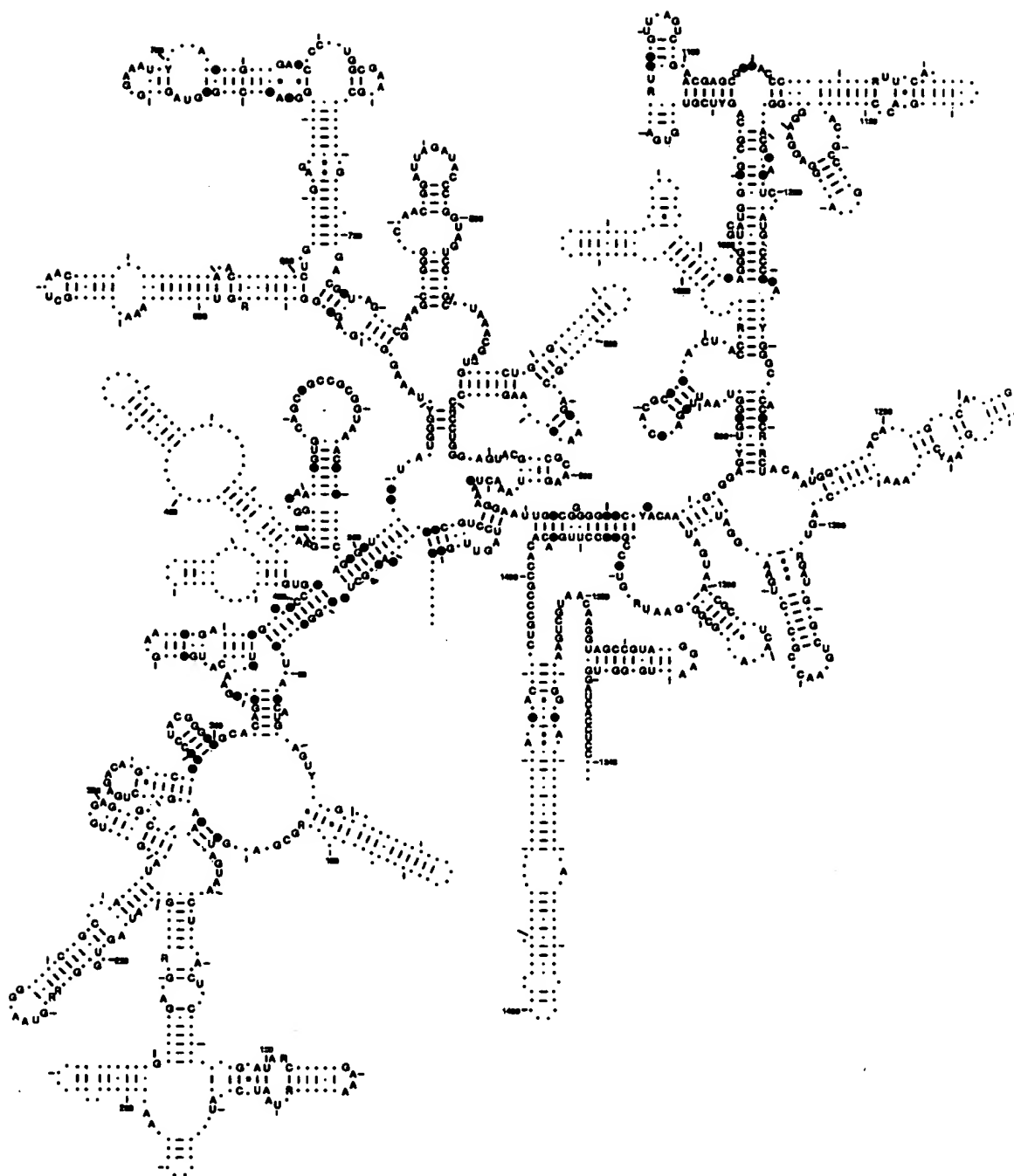


FIG. 7. Location of the signature positions that distinguish eubacteria from archaebacteria (Table 1) on the 16S rRNA secondary structure. The underlying secondary structure is that for *Escherichia coli* (260). The signature positions are indicated by filled circles. Positions whose composition is highly conserved (i.e., the same in over 90% of sequences and oligonucleotide catalogs) between eubacteria and archaebacteria are also indicated (as an aid in orientation); all others (in the *E. coli* sequence) have been replaced by dots. Tick marks show every 10th position in the (*E. coli*) sequences, every 50th being numbered.

now understood are listed in Table 2. Table 3 identifies them by a sequence signature. This key now includes mainly positions covered by the oligonucleotide catalogs, but will ultimately be extended to all positions in the 16S (and 23S) rRNA.

In the following discussion each of the known eubacterial phyla (divisions) will be defined genealogically and its phenotype will be briefly described. As was the case with the kingdoms themselves, three criteria will be used whenever

TABLE 2. Eubacterial phyla and their subdivisions<sup>a</sup>

Purple bacteria
α subdivision
Purple non-sulfur bacteria, rhizobacteria, agrobacteria, rickettsiae, <i>Nitrobacter</i>
β subdivision
<i>Rhodocyclus</i> , (some) <i>Thiobacillus</i> , <i>Alcaligenes</i> , <i>Spirillum</i> , <i>Nitrosovibrio</i>
γ subdivision
Enterics, fluorescent pseudomonads, purple sulfur bacteria, <i>Legionella</i> , (some) <i>Beggiatoa</i>
δ subdivision
Sulfur and sulfate reducers ( <i>Desulfovibrio</i> ), myxobacteria, bedellovibrions
Gram-positive eubacteria
A. High-G+C species
<i>Actinomyces</i> , <i>Streptomyces</i> , <i>Arthrobacter</i> , <i>Micrococcus</i> , <i>Bifidobacterium</i>
B. Low-G+C species
<i>Clostridium</i> , <i>Peptococcus</i> , <i>Bacillus</i> , mycoplasmas
C. Photosynthetic species
<i>Hellobacterium</i>
D. Species with gram-negative walls
<i>Megasphaera</i> , <i>Sporomusa</i>
Cyanobacteria and chloroplasts
<i>Aphanocapsa</i> , <i>Oscillatoria</i> , <i>Nostoc</i> , <i>Synechococcus</i> , <i>Gloeobacter</i> , <i>Prochloron</i>
Spirochetes and relatives
A. Spirochetes
<i>Spirochaeta</i> , <i>Treponema</i> , <i>Borrelia</i>
B. Leptospiras
<i>Leptospira</i> , <i>Leptonema</i>
Green sulfur bacteria
<i>Chlorobium</i> , <i>Chloroherpeton</i>
Bacteroides, flavobacteria, and relatives
A. Bacteroides
<i>Bacteroides</i> , <i>Fusobacterium</i>
B. Flavobacterium group
<i>Flavobacterium</i> , <i>Cytophaga</i> , <i>Saprospira</i> , <i>Flexibacter</i>
Planctomyces and relatives
A. Planctomyces group
<i>Planctomyces</i> , <i>Pasteuria</i>
B. Thermophiles
<i>Isocystis pallida</i>
Chlamydiae
<i>Chlamydia psittaci</i> , <i>C. trachomatis</i>
Radioresistant micrococci and relatives
A. Deinococcus group
<i>Deinococcus radiodurans</i>
B. Thermophiles
<i>Thermus aquaticus</i>
Green non-sulfur bacteria and relatives
A. Chloroflexus group
<i>Chloroflexus</i> , <i>Herpetosiphon</i>
B. Thermomicrobium group
<i>Thermomicrobium roseum</i>

<sup>a</sup> Showing representative examples. See appropriate sections in text for references.

possible to define a phylum and its subdivisions: (i) coherence of the unit by sequence distance analysis (or cluster analysis of  $S_{AB}$  values [55]); (ii) definition of the unit by sequence signature; and (iii) characterization of the unit in terms of higher-order structural features of 16S rRNA.

### Purple Bacteria

What for want of a better name have been called the "purple bacteria" contain most, but not all, of the traditional gram-negative bacteria (50, 51, 56, 64, 163, 254, 266-269). However, the arrangement of classically defined families, genera, and even species within this phylum is a jumbled one. Photosynthetic species group with nonphotosynthetic ones; heterotrophs are paired with chemolithotrophs; anaerobes are paired with aerobes, etc. Because the purple photosynthetic phenotype is distributed more or less throughout the group, and because photosynthesis is complex enough that its having arisen more than once is unlikely, the ancestral phenotype of the phylum is undoubtedly (purple) photosynthetic, which justifies the name purple bacteria. Photosynthetic capacity has been lost many times in this phylum, resulting in various nonphotosynthetic sublines. (The alternative explanation, that photosynthetic capacity was genetically transferred among species, is not supported by the evidence [258].)

The purple bacteria fall rather naturally into four subdivisions, which, awaiting appropriate formal nomenclature, are designated α, β, γ, and δ. Figure 8 is a distance matrix tree for the purple bacteria, based upon five representative sequences in each subdivision. The corresponding tree given by maximum parsimony analysis (not shown) agrees with the branching order shown in Fig. 8, except for some details of branching in the δ subdivision. Table 4 distinguishes the four subdivisions by sequence signature. The purple bacteria as a whole, however, make up one of the few phyla that cannot be defined by a simple signature (see Table 3), although parsimony analysis as well as sequence distance treeing readily define the group (267).

Two helices in 16S rRNA, positions 184 to 193 and 198 to 219 (*E. coli* numbers) (260), help to define and distinguish the four subdivisions; see Fig. 9. While the structure of these helices can vary drastically between subdivisions, within a given subdivision each remains constant to a first approximation. The first of the two helices, i.e., positions 184 to 193, contains only 3 base pairs in α, β, and γ subdivision sequences, but about 10 base pairs in the δ subdivision sequences. The short form of the helix is rare. So far it has been found outside the purple bacteria in only two other phyla, the cyanobacteria (224) and the planctomyces (H. Oyaizu, unpublished data). In all other cases, including the archaeobacteria (72, 96, 124, 126, 158), a much longer version of the helix occurs, which therefore is likely to be its ancestral form.

The second helix, positions 198 to 219, contains approximately 8 base pairs in all, except the α-subdivision sequences, wherein 2 base pairs only make up the stalk (260, 274). However, for this helix the short version is common among eubacteria and even occurs in archaeobacteria, while the longer version is rare (among eubacteria). The long version has a common structure in most sequences from the β and γ subdivisions, but the characteristic δ-subdivision (long) version differs from this structure in detail (see Fig. 9).

The (near) constancy of structure for each helix within a given subdivision can be inferred from oligonucleotide catalog data. For example, 80% of α-purple bacterial catalogs

TABLE 3. 16S rRNA sequence signature for the eubacterial phyla and their subdivisions<sup>a</sup>

Position	Consensus composition	Purple bacteria				Gram-positive bacteria		Cyanobacteria	Green sulfur bacteria	Spirochetes		Bacteroides		Planctomyces	Deinococcus	Green non-sulfur bacteria
		$\alpha$	$\beta$	$\gamma$	$\delta$	Lo	High			Spirochetes	Leptospiras	Bacteroides	Flavobacteria <sup>b</sup>			
47	C	• <sup>b</sup>	•	•	•	•	•	•	•	U <sup>c</sup>	U	•	•	G	•	•
48	Y	•	•	•	•	•	•	•	•	•	•	•	•	A	•	•
50	A	•	U	•	•	•	•	•	•	U	U	•	•	U	•	•
52	Y	•	•	•	•	•	•	•	•	A	A	•	•	G	G	•
53	A	•	•	•	•	•	•	•	•	G	•	•	•	G	•	G
353	A	•	•	•	•	•	•	•	•	•	•	•	•	U	•	•
570	G	•	•	•	•	•	•	•	•	•	•	U	U	U	•	•
698	G	U	•	•	Y	•	•	•	•	•	•	•	•	•	•	•
812	G	•	C	• <sup>d</sup>	•	•	•	•	•	•	C	•	• <sup>e</sup>	•	C	•
906	G	Ag	Ag	Ag	Ag	• <sup>a</sup>	A	•	•	•	•	•	•	•	•	A
933	G	•	•	•	•	•	•	•	• <sup>a</sup>	•	•	•	•	A	•	•
955	U	•	Au	•	•	•	AC	•	•	•	•	•	•	C	•	•
976	G	•	A	•	•	•	•	•	•	• <sup>a</sup>	A	•	•	•	•	•
983:1	— <sup>f</sup>	—	—	—	—	—	—	—	—	—	—	—	—	U	—	—
995	C	•	•	•	•	•	•	•	(A) <sup>g</sup>	•	•	A	A	•	•	•
1109	C	•	•	•	•	•	•	•	•	•	•	•	•	Ac	•	•
1198	G	•	•	•	•	Ag	Ag	•	•	• <sup>a</sup>	• <sup>a</sup>	A	Ag	•	•	•
1202	U	•	•	•	•	•	•	• <sup>g</sup>	•	•	•	G	•	•	g <sup>h</sup>	G
1207	G	•	•	•	•	C	C	C	•	•	•	•	•	•	•	•
1208	C	•	•	•	•	•	•	•	•	•	•	•	•	•	u <sup>h</sup>	•
1224	U	•	•	•	•	• <sup>c</sup>	• <sup>a</sup>	• <sup>g</sup>	YA	• <sup>c</sup>	UA	• <sup>g</sup>	• <sup>c</sup>	•	•	G
1229	A	•	•	•	•	•	Ga	•	•	•	•	•	•	•	•	•
1233	G	•	Ca	•	• <sup>a</sup>	a	•	A	•	• <sup>a</sup>	•	•	•	• <sup>c</sup>	•	•
1234	C	•	A	•	• <sup>a</sup>	•	•	•	A	• <sup>a</sup>	•	U	•	•	•	•
1384	C	•	•	•	•	•	•	•	• <sup>u</sup>	•	•	•	•	U	•	•
1410	A	•	•	•	•	•	G	•	G	•	•	G	G	G	•	G
1415	G	•	•	•	•	•	•	•	•	C	C	•	•	•	•	•
1520	G	Cg	• <sup>c</sup>	Cg	C	• <sup>c</sup>	•	•	•	• <sup>a</sup>	Cg	•	•	(C)	•	•
1532	U	•	•	•	•	•	•	•	•	•	•	Au	Au	•	•	•

<sup>a</sup> Abbreviations not explained are obvious from text. All positions are based upon oligonucleotide data (266) except 353 (233).<sup>b</sup> Same composition as consensus (•).<sup>c</sup> Composition upper case—major base; if no other specified, then it accounts for >90% of assayable cases.<sup>d</sup> Composition lower case—minor occurrence base; found in <15% of assayable cases (or in only one species for groups containing seven or less species).<sup>e</sup> —, No nucleotide at this position.<sup>f</sup> Composition in parentheses—based upon one example only.

contain the octanucleotide (G)AUUUAUCG, which entirely covers the version of the second helix (positions 198 to 219) found in this subdivision (266, 267). (G)AYCUUCG, which forms part of this helix in the  $\beta$  and  $\gamma$  subdivisions, is found in 44% of  $\gamma$  catalogs and 13% of those from the  $\beta$  subdivision (267–269). (Although all known 16S rRNA sequences from  $\beta$  species have the structure for this helix shown in Fig. 9, the loop often contains a G residue, which breaks up the [T<sub>1</sub> ribonuclease] oligonucleotide otherwise characteristic of the structure [unpublished analysis].) Oligonucleotides of the form (G)CCUCU . . . , seen in the  $\delta$ -purple bacterial version of this helix, occur in 78% of  $\delta$ -subdivision catalogs (163).

With regard to the first helix, positions 184 to 193, oligonucleotides representing the structure can be identified in only 35 and 26% of cataloged species, respectively, from the  $\beta$  and  $\gamma$  subdivisions (unpublished analysis). However, every one of the sequenced 16S rRNAs from these two subdivisions show the characteristic 3-base pair structure (unpublished analysis). Based upon oligonucleotide and sequence evidence, at least 83% of  $\alpha$ -purple bacteria must also possess the same form of this helix (unpublished analysis).

Figure 8 shows the  $\beta$  and  $\gamma$  subdivisions to be relatively closely related to one another, a fact that can also be demonstrated with other types of analysis (163); see Table 4. The common  $\beta$ - $\gamma$  lineage and the  $\alpha$ - and  $\delta$ -subdivision lineages appear to have split from one another in such rapid succession that their branching order cannot be resolved by the evidence that now exists. However, the higher-order structural evidence, specifically the (derived) form of the helix covering positions 184 to 193 (Fig. 9), would suggest that the  $\alpha$ ,  $\beta$ , and  $\gamma$  subdivisions form a grouping that excludes the  $\delta$  subdivision.

Table 5 through 8 lists representative genera and species in each of the four subdivisions in a rough phylogenetic arrangement. Three of the four subdivisions contain photosynthetic species: the  $\alpha$ -purple bacteria seems predominantly photosynthetic, and the  $\beta$  subdivision shows photosynthesis in its two main subgroups, while in the  $\gamma$  subdivision photosynthesis appears to be confined to one of its three main subgroups.

**$\alpha$ -Purple bacteria.** The intimate juxtaposition of photosynthetic and nonphotosynthetic species in the  $\alpha$  subdivision

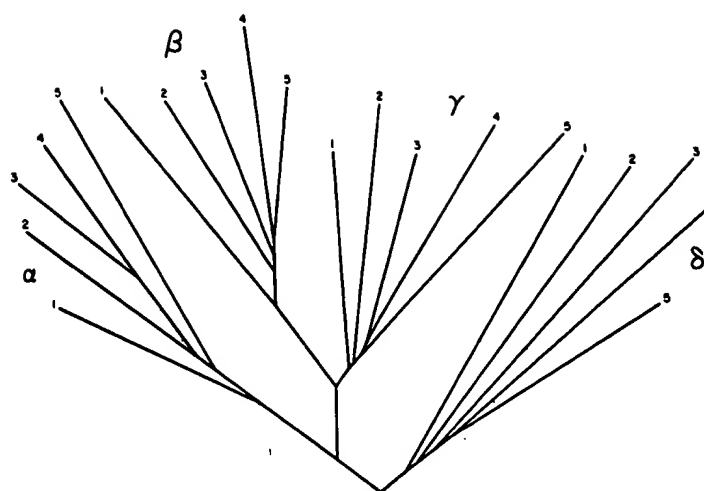


FIG. 8. Phylogenetic tree for the purple bacteria based upon 16S rRNA sequences. The tree was constructed (36) from an evolutionary distance matrix (99), generated from an alignment (260) of five representative sequences from each of the four purple bacterial subdivisions. (Only positions represented in all sequences were used in the calculation.) The root was determined by using several eubacterial outgroup sequences. The sequences used were as follows.  $\alpha$  subdivision: 1, *Rhodospirillum rubrum* (Woese et al., unpublished data); 2, *Agrobacterium tumefaciens* (275); 3, *Rhodopseudomonas palustris* (Woese et al., unpublished data); 4, *Rhodopseudomonas acidophila* (Woese et al., unpublished data); 5, *Rhodobacter capsulatum* (Woese et al., unpublished data).  $\beta$  subdivision: 1, *Neisseria gonorrhoeae* (Woese et al., unpublished data); 2, *Spirillum volutans* (Woese et al., unpublished data); 3, *Nitrosolobus multififormis* (Woese et al., unpublished data); 4, *Rhodocyclus gelatinosa* (Woese et al., unpublished data); 5, *Rhodocyclus pupureus* (Woese et al., unpublished data).  $\gamma$  subdivision: 1, *Chromatium vinosum* (Woese et al., unpublished data); 2, *Legionella pneumophila* (Woese et al., unpublished data); 3, *Pseudomonas aeruginosa* (Woese et al., unpublished data); 4, *Acinetobacter calcoaceticus* (Woese et al., unpublished data); 5, *Escherichia coli* (19).  $\delta$  subdivision: 1, *Myxococcus xanthus* (163); 2, *Desulfovibrio desulfuricans* (163); 3, *Bdellovibrio stolpii* (Woese et al., unpublished data); 4, *Desulfotobacter postgatei* (Woese et al., unpublished data); 5, *Desulfuromonas acetoxidans* (Woese et al., unpublished data). Outgroup sequences: *Bacillus subtilis* (68) and *Thermotoga maritima* (1).

suggests a more or less continual evolution of the latter from the former (267). Aerobic metabolism also appears to have arisen a number of times in this subdivision alone (267). The close association of species reducing and oxidizing nitrogen compounds, e.g., *Rhodopseudomonas palustris* and *Nitrobacter winogradskii* (267), suggests some sort of evolutionary connection between the two metabolisms. The metabolic richness and diversity of species evolving from purple photosynthetic ancestry in this subdivision are remarkable.

The  $\alpha$  subdivision is also of general biological interest because certain of its members have interesting associations with various eucaryotes. The rhizobacteria (essential for nitrogen fixation in legumes), the agrobacteria (pathogenic for plants) (267, 274), and the rickettsias (intracellular pathogens of animals) (235) form a tight cluster within subgroup  $\alpha$ -2. Sequence differences in 16S rRNA among members of this cluster are under 7% (235). These particular species have in common the tendency to form intimate, if not intracellular, associations with eucaryotic cells. It is then no great surprise to learn that the endosymbiont that gave rise to (most, if not all) eucaryotic mitochondria was itself a member of the  $\alpha$  subdivision (274).

**$\beta$ -Purple bacteria.** Most of the characterized  $\beta$ -purple species fall into two main subgroups. However, poorly defined deeper branchings are also evident, represented, for example, by *Spirillum* and *Neisseria* (269). This subdivision is a potpourri of classical genera (Table 6), some of which are not even phylogenetically coherent within the subdivision (269). The  $\beta$ -photosynthetic species, recently reclassified in the genus *Rhodocyclus* (92), are quite distinct from other purple nonsulfur bacteria, i.e., those residing in the  $\alpha$  subdivision. Over and above their rRNA sequence differ-

ences,  $\beta$  species differ from their  $\alpha$  counterparts in cytochrome *c* type;  $\beta$  cytochromes are of the small-subunit type, while cytochromes from the  $\alpha$  subdivision are of the medium or large type (37). Moreover, photoreaction centers in photosynthetic  $\beta$  species have a structure distinct from that seen among  $\alpha$  species (27).

**$\gamma$ -Purple bacteria.** The  $\gamma$ -purple bacteria (Table 7), the most extensively characterized of the four subdivisions of purple bacteria, is again a mixture of phenotypes (268): photosynthetic with nonphotosynthetic, aerobic with anaerobic, heterotrophic with chemolithotrophic, etc. Oligonucleotide catalog analysis divided the  $\gamma$ -purple bacteria into three main subgroups: one containing mainly photosynthetic species of the purple sulfur type, e.g., *Chromatium* (50, 198); a second known to contain only species associated with Legionnaires disease (132); and a third that is a mixture of nonphotosynthetic genera from the enterics, vibrios, oceanospirilla, the fluorescent pseudomonads and relatives, and others (268). As additional complete 16S rRNA sequences become available, it begins to look as though the  $\beta$  subdivision may ultimately be shown to be a subgroup within the  $\gamma$  subdivision (albeit deeply branching). In any case its close association with the  $\gamma$  subdivision is surprising given that the  $\beta$ -purple bacteria have such a distinctive signature; see Table 4 (266).

**$\delta$ -Purple bacteria.** The  $\delta$ -purple bacteria (Table 8) subdivision harbors three disparate phenotypes: the sulfur and sulfate-reducing eubacteria, the myxobacteria and relatives, and the bdellovibrios (163). At the present writing the phylogenetic detail within the subdivision is not clear. The myxobacteria and relatives indeed form a coherent grouping, a clade (130). The bdellovibrios probably form one as well,



TABLE 5. Characterized species of  $\alpha$ -purple bacteria (64, 269)<sup>a</sup>

Subgroup $\alpha$ -1
<i>Rhodospirillum rubrum</i>
<i>R. photometricum</i>
<i>R. molischianum</i>
<i>R. fulvum</i>
<i>Rhodopseudomonas globiformis</i>
<i>Aquaspirillum itersonii</i>
<i>Azospirillum brasilense</i>
Subgroup $\alpha$ -2
<i>Rhodomicrobium vannielii</i>
<i>Rhodopseudomonas viridis</i>
<i>Rhodopseudomonas palustris</i>
<i>Nitrobacter winogradskyi</i>
<i>Rhizobium leguminosarum</i>
<i>Agrobacterium tumefaciens</i> <sup>a</sup>
<i>Rochalimaea quintana</i>
<i>Rhodopseudomonas acidophila</i>
<i>Pseudomonas diminuta</i>
<i>Phenylobacterium immobile</i>
Subgroup $\alpha$ -3
<i>Rhodobacter capsulatum</i>
<i>Rhodobacter sphaeroides</i>
<i>Paracoccus denitrificans</i>
Manganese oxidizers (2 strains)
Subgroup $\alpha$ -4
<i>Erythrobacter longus</i>

<sup>a</sup> Indentation indicates specific relationship; for example, *Rhizobium leguminosarum*, *Agrobacterium tumefaciens*, and *Rochalimaea quintana* are specific relatives of one another, to the exclusion of the other species in their subgroup.

TABLE 6. Characterized species of  $\beta$ -purple bacteria (64, 269)<sup>a</sup>

Subgroup $\beta$ -1
<i>Rhodocyclus gelatinosa</i>
<i>Sphaerotilus natans</i>
<i>Pseudomonas testosteroni</i>
<i>P. acidovorans</i>
<i>Aquaspirillum gracile</i>
<i>A. aquaticum</i>
<i>Comamonas terrigena</i>
<i>Thiobacillus intermedius</i>
Subgroup $\beta$ -2
<i>Rhodocyclus tenue</i>
<i>R. purpureus</i>
<i>Aquaspirillum dispar</i>
<i>A. serpens</i>
<i>A. bengal</i>
<i>Chromobacterium violaceum</i>
<i>Chromobacterium lividum</i>
<i>Alcaligenes faecalis</i>
<i>Alcaligenes eutrophus</i>
<i>Pseudomonas cepacia</i>
<i>Thiobacillus denitrificans</i>
<i>Vitreoscilla stercoraria</i>
Subgroup $\beta$ -3 <sup>b</sup>
<i>Spirillum volutans</i>
<i>Nitrosomonas europaea</i>
<i>Nitrosococcus mobilis</i>
<i>Nitrosolobus multififormis</i>
<i>Nitrosovibrio tenuis</i>
<i>Nitrosospora</i> sp.
<i>Neisseria gonorrhoeae</i>

<sup>a</sup> Indentation indicates specific relationship.

<sup>b</sup> Paraphyletic group, defined only by exclusion from subgroups 1 and 2.

TABLE 7. Characterized genera or species of  $\gamma$ -purple bacteria (50, 64, 132, 268)<sup>a</sup>

Subgroup $\gamma$ -1
<i>Chromatium</i>
<i>Thiocapsa</i>
<i>Thiocystis</i>
<i>Thiodictyon</i>
<i>Thiospirillum</i>
<i>Lamprocystis</i>
<i>Nitrosococcus oceanus</i>
<i>Ectothiorhodospira</i>
Subgroup $\gamma$ -2
<i>Legionella</i>
Subgroup $\gamma$ -3
Fluorescent pseudomonads <sup>b</sup>
<i>Xanthomonas</i>
<i>Lysobacter</i>
<i>Acinetobacter</i>
<i>Oceanospirillum</i>
<i>Alcaligenes putrefaciens</i>
<i>Pasteurella multocida</i>
<i>Aeromonas hydrophila</i>
" <i>Bacteroides</i> " <i>amylophilus</i>
Enterics, vibrios, and photobacteria
<i>Halomonas elongata</i>
" <i>Flavobacterium</i> " <i>helmeophilum</i>
<i>Leucothrix mucor</i>
<i>Beggiatoa leptomitiformis</i>

<sup>a</sup> Indentation indicates specific relationship.

<sup>b</sup> Includes *Pseudomonas aeruginosa*, *P. alcaligenes*, *P. fluorescens*, *P. putida*, *P. stutzeri*, *P. syringae*, *P. pseudoalcaligenes*, and *Serpens flexibilis*.

TABLE 8. Characterized genera and species of  $\delta$ -purple bacteria (51, 83, 130, 163)<sup>a</sup>

Myxococcus group
<i>Myxococcus</i>
<i>Cystobacter fuscus</i>
<i>Stigmatella aurantiaca</i>
<i>Sorangium cellulosum</i>
<i>Nannocystis exedens</i>
Bdellovibrio group
<i>Bdellovibrio stolpii</i>
<i>Bdellovibrio starii</i>
<i>Bdellovibrio bacteriovorus</i>
Sulfur and sulfate reducers
<i>Desulfovibrio desulfuricans</i>
<i>Desulfuromonas</i>
<i>Desulfotobacter postgatei</i>
<i>Desulfosarcina variabilis</i>
<i>Desulfonema limicola</i>
<i>Desulfobulbus propionicus</i> ??

<sup>a</sup> Indentation indicates specific relationship.

but the fact that *Bdellovibrio bacteriovorus* is relatively rapidly evolving makes its exact placement uncertain (83). Whether the myxobacteria and relatives or the bdellovibrios (or both) arise outside of the group defined by the sulfur and sulfate reducers remains unresolved (unpublished analysis). In any case it would appear that the myxobacteria and bdellovibrios represent aerobic adaptations of some ancestral anaerobic sulfur-metabolizing phenotype.

#### Gram-Positive Eubacteria

Cell wall type distinguishes the gram-positive eubacteria from the others (56, 183). However, as mentioned above,



TABLE 9. Characterized genera and species of gram-positive bacteria\*

High-G+C subdivision	Low-G+C subdivision
<i>Bifidobacterium</i>	<i>Bacillus</i>
<i>Propionibacterium</i>	<i>Planococcus</i>
<i>Actinomyces</i>	<i>Sporolactobacillus</i>
<i>Arthrobacter</i>	<i>Sporosarcina</i>
<i>Micrococcus</i>	<i>Thermoactinomyces</i>
<i>Dermatophilus</i>	<i>Staphylococcus</i>
<i>Cellulomonas</i>	<i>Lactobacillus</i>
<i>Derskovia</i>	<i>Pediococcus</i>
<i>Nocardia cellulans</i>	<i>Leuconostoc</i>
<i>Microbacterium</i>	<i>Streptococcus</i>
<i>Corynebacterium</i> (plant associated)	<i>Mycoplasma</i>
<i>Brevibacterium linens</i>	<i>Acholeplasma</i>
<i>Streptomyces</i>	<i>Spiroplasma</i>
<i>Kitasatoa</i>	<i>Anaeroplasm</i>
<i>Chainia</i>	<i>Clostridium innocuum</i>
<i>Microellorospira</i>	<i>Erysipelothrix</i>
<i>Streptovercillium</i>	<i>Clostridium pasteurianum</i>
<i>Actinomadura</i>	<i>C. butyricum</i>
<i>Streptosporangium</i>	<i>C. scatologenes</i>
<i>Thermomonospora</i>	<i>Sarcina ventriculae</i>
<i>Mycobacterium</i>	<i>Clostridium oroticum</i>
<i>Nocardia</i>	<i>C. indolis</i>
<i>Brevibacterium ketoglutamicum</i>	<i>C. aminovalericum</i>
<i>Corynebacterium</i>	<i>Butyrivibrio fibrosolvens</i>
<i>Geodermatophilus</i>	<i>Clostridium lituseburense</i>
<i>Frankia</i>	<i>Eubacterium tenue</i>
<i>Dactylosporangium</i>	<i>Peptostreptococcus</i>
<i>Ampurariella</i>	<i>C. aceticum</i>
<i>Actinoplanes</i>	<i>C. acidurici</i>
<i>Micromonospora</i>	<i>C. purinolyticum</i>
<i>Arthrobacter simplex</i>	<i>Clostridium barkeri</i>
	<i>Eubacterium limosum</i>
	<i>Acetobacterium woodii</i>
	<i>Clostridium thermosaccharolyticum</i>
	<i>C. thermoaceticum</i>
	<i>Acetogenium</i>
	<i>Thermoanaerobium</i>
	<i>Peptococcus</i>
	<i>Ruminococcus</i>
Photosynthetic subdivision	
<i>Heliobacterium</i>	
Species with gram-negative walls	
<i>Megasphaera</i>	
<i>Selenomonas</i>	
<i>Sporomusa</i>	

\* Approximate phylogenetic clustering suggested by indentation. References 56, 131, 199-201, 203-207, 220, 221, 255, 262, 264.

there are several important exceptions whose walls are not gram positive. The phylum appears to consist of four subdivisions, only two of which are well characterized. These two are readily distinguished on the basis of DNA composition. The one includes species whose DNAs all contain more than 55% guanine plus cytosine (G+C); the other is made up of species whose DNAs contain <50% (56, 201, 204). The recently discovered phototroph *Heliobacterium chlorum* (58, 60, 255) is the only characterized representative of the third subdivision, while the genera *Megasphaera*, *Selenomonas*, and *Sporomusa* constitute the fourth (203). Members of the third and fourth subdivisions do not have gram-positive cell walls (60, 203).

Species in the high-G+C gram-positive subdivision conform to a general actinomycete phenotype: they tend to be pleomorphic, form branched filaments, etc. (200, 201, 204-206). Most are aerobic, with the exception of the deeper branches, e.g., the bifidobacteria. The group as a whole is not particularly deep; by oligonucleotide catalog measure all high-G+C gram-positive bacteria are no further from each other than are *Bacillus* species from those of *Lactobacillus*, for example. The lowest  $S_{AB}$  values in the group correspond roughly to sequence similarities in the range of 85% (56, 201, 204). The group, therefore, would seem not to be a particularly ancient one.

Species in the low-G+C gram-positive subdivision conform by and large to a clostridial phenotype. They tend to be

anaerobic, rod shaped, endosporeforming, etc., although a number have lost one or more of these characteristics. In contrast to the high-G+C group, the low-G+C gram-positive bacteria form a phylogenetically deep, and therefore presumably ancient, cluster (56).

Table 9 lists some of the characterized gram-positive genera and species in a crude phylogenetic ordering.

The gram-positive rRNAs sequenced so far are not broadly representative enough to permit the construction of a phylogenetic tree for the entire phylum. However, the phylum is easily defined by cluster analysis of  $S_{AB}$  values based upon oligonucleotides (56, 206). Although few in number, the signature positions characterizing the gram-positive eubacteria (Table 3) are strong ones. Note in particular the presence of a C residue at position 1207, which occurs in all cataloged gram-positive bacteria (roughly 150 species) and in cyanobacteria, but nowhere else among eubacteria (266). The A residue at position 1198 is present in roughly 75% of gram-positive species, but occurs elsewhere only in the bacteroides phylum and occasionally among the spirochetes (266). An absolute requirement for an A residue at position 513 also characterizes the gram-positive bacteria, but few other phyla.

Higher-order structure in 16S rRNA also helps to identify gram-positive bacteria. Two characteristic adjacent A-G pairs, positions 1425-6 to 1474-5, occur in the penultimate helix of almost all gram-positive 16S rRNAs sequenced to

TABLE 10. Sequence signature distinguishing gram-positive subdivisions<sup>a</sup>

Position <sup>b</sup>	Subdivision	Composition (%) <sup>c</sup>				Consensus <sup>d</sup>
		A	G	C	U	
168	Low G+C	5	81	0	0	G
	High G+C	2	20	0	70	
	Sporomusa group	0	100	0	0	
	<i>H. chlorum</i>	—	+	—	—	
906	Low G+C	7	92	0	0	G
	High G+C	100	0	0	0	
	Sporomusa group	0	100	0	0	
	<i>H. chlorum</i>	—	+	—	—	
955	Low G+C	0	0	1	99	U
	High G+C	74	0	25	1?	
	Sporomusa group	0	0	0	100	
	<i>H. chlorum</i>	—	—	—	+	
998	Low G+C	2	5	88	3	N
	High G+C	36	33	16	1	
	Sporomusa group	0	0	0	100	
	<i>H. chlorum</i>	—	—	+	—	
1116	Low G+C	0	0	1	95	U
	High G+C	0	0	54	46	
	Sporomusa group	0	0	0	100	
	<i>H. chlorum</i>	—	—	—	+	
1167	Low G+C	62	?	2	1	A
	High G+C	0	0	2	98	
	Sporomusa group	75	0	0	0	
	<i>H. chlorum</i>	+	—	—	—	
1410	Low G+C	94	<6	0	0	A
	High G+C	0	>46	0	0	
	Sporomusa group	75	0?	0	0	
	<i>H. chlorum</i>	+	—	—	—	

<sup>a</sup> References: low-G+C subdivision (131, 199, 203, 220, 221); high-G+C subdivision (200, 201, 204, 205, 207); sporomusa group (203); *Hellobacterium chlorum* (255).

<sup>b</sup> Position in sequence.

<sup>c</sup> Composition of position (percentage of oligonucleotide catalogs having an oligonucleotide covering the position which shows the indicated composition); failure to find an oligonucleotide covering a position in some catalogs causes percentage compositions to sum to less than 100%.

<sup>d</sup> Composition characteristic of position in the (vast) majority of other eubacterial phyla.

date (260; unpublished data); see Fig. 10. This arrangement has yet to be seen outside this group. (We will encounter G-A pairs at this location in another phylum, however.) Oligonucleotide catalog data indicate that these adjacent A-G pairs are common among gram-positive bacteria. The general sequence (G)YAAAYACCC (which includes the two adjacent A's of the A-G pairs in question) is found in 85% of catalogs from the high-G+C group and in 67% of those from the low-G+C group, but occurs nowhere else among eubacteria (266).

The sequence signature of Table 10 distinguishes between the two main gram-positive subdivisions. In all cases the composition characteristic of the low-G+C subdivision is that found in most other phyla; it is undoubtedly, therefore, the ancestral composition. This would suggest that the high-G+C gram-positive lineage has been subject to rapid evolution; see discussion below.

Genealogical substructure within the two major subdivisions can be seen (56, 200, 201, 204–206), although it will not be systematically detailed here. In the high-G+C subdivision, the deepest branchings are defined by anaerobic species, the bifidobacteria and the propionibacteria. The low-G+C subdivision shows at least five major branches, most of which contain clostridia.

One subline in the low-G+C subdivision has given rise to four groups of particular interest: *Bacillus*, *Lactobacillus*, *Streptococcus*, and the mycoplasmas and their relatives (56, 262). The subline is an interesting evolutionary study, for its evolution in a way parallels the development of aerobic conditions on the planet. *Bacillus* species are basically aerobic, though a few also grow well anaerobically. *Lactobacillus*, *Streptococcus*, and the mycoplasmas are basically anaerobic, but tolerate and in some cases even utilize a little oxygen. Their clostridial relatives are true anaerobes. The evolutionary radiation that spawned the various groups might well have occurred during the microaerobic period in earth history (see below), with *Bacillus* then becoming fully adapted to an oxygen atmosphere. (The exact order of branching among these groups will not be known until sequences representing all four have been determined.) The mycoplasma group will be considered in detail below.

#### Cyanobacteria

The cyanobacteria, the classical blue-green algae, are a group of procaryotes defined by the common possession of chlorophyll *a*. They form a phylogenetically coherent unit (12, 40, 56) that contains no known nonphotosynthetic

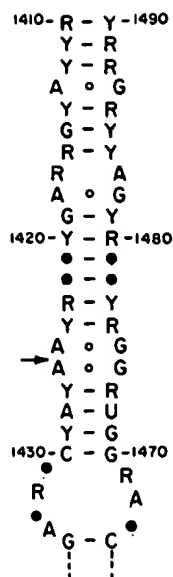


FIG. 10. Higher-order structure for the gram-positive bacteria in the region of 16S rRNA between positions 1410 and 1480 (68, 93, 163). Composition of a position is given when it is invariant or highly conserved within the gram-positive bacteria, but it is shown as a dot otherwise. Base pairs are indicated by connecting lines; A-G pairs are shown by open circles. Arrow shows adjacent A-G pairs discussed in text.

representatives. However, the phylum does include *Prochloron* (191), an organism that, like green plant chloroplasts, possesses chlorophyll *b* as well as chlorophyll *a* (127). Green algal chloroplasts trace their ancestry to this phylum, as expected (11, 41, 67, 189, 223). Analysis of oligonucleotide catalog data did not resolve the questions of whether the chloroplasts (whose rRNAs appear somewhat rapidly evolving) arose from within or from just outside of the cluster defined by extant cyanobacteria and whether these organelles are related to *Prochloron* to the exclusion of the cyanobacteria. Although about 25 complete sequences of cyanobacterial 16S rRNAs have now been determined (but as yet not that from *Prochloron*), these questions remain unresolved as of this writing (S. J. Giovannoni, S. Turner, G. J. Olsen, D. J. Lane, and N. R. Pace, manuscript in preparation). The precise origin of the brown algal chloroplasts and certain others also has yet to be determined. However, the rRNAs of the chloroplasts from the red algae (10) and *Euglena* sp. (276) are known to be closely related to those of cyanobacteria.

The sequence signature for the cyanobacteria and chloroplasts (Table 3) is small but significant. Position 799, a G residue in all but 1 to 2% of other eubacterial catalogs, is U or A in all cyanobacterial sequences or catalogs, including *Prochloron* (191); G's are encountered in some chloroplast sequences, however (276). Position 1207, a G residue in all other eubacterial catalogs, is always C in the gram-positive bacteria and cyanobacteria (including *Prochloron* and the chloroplasts), except for one cyanobacterial sequence (191; Giovannoni et al., in preparation). Very few eubacteria show A at position 1233 (266); however, all cyanobacteria (except one) and *Prochloron* do (191; Giovannoni et al., in prepara-

tion). (Most chloroplast examples do not [10, 41, 67, 189, 223, 276].)

#### Green Sulfur Bacteria

The four cataloged species of the green photosynthetic bacteria covering the genera *Chlorobium* and *Chlorohelium* form a relatively tight phylogenetic unit, especially in view of the fact that their phenotype is generally considered a very primitive one (63, 266). Since the characterized species seem representative of the known spectrum of green sulfur bacteria, the question of why the group is relatively shallow becomes nontrivial. As will be seen below, the green sulfur bacteria are not related to the other green type of photosynthetic bacteria, the so-called green non-sulfur bacteria (63).

The signature characteristic of the phylum (Table 3) contains positions 995, 1234, and 1410. The one sequence now available for the group (W. G. Weisburg, Ph.D. thesis, University of Illinois, Urbana, 1986) suggests that additional strong signature positions will appear when more sequences from the group are known, e.g., the lack of a base in the vicinity of position 1167 and the insertion of a base after position 1174 (Weisburg, Ph.D. thesis).

#### Spirochetes

Spirochetes are one of the few groupings correctly identified by classical (morphological) criteria (167). Their common spiral shape and axially coiled fibrils, lying between inner and outer cell envelopes, are strikingly characteristic (24). Table 11 shows representative species whose rRNAs have been characterized, in rough phylogenetic arrangement. The sequence signature for the group shown in Table 3 is quite distinctive. For example, the U residue at position 47 found in all species from this group occurs nowhere else among eubacteria (266; unpublished analysis). The same holds for the A residue at position 52, and the C residue at position 1415, while universal among spirochetes, is otherwise extremely rare among eubacteria (266).

Two clearly separated subdivisions exist within the phylum: one composed of the leptospiras and the other containing spirochetes, treponemes, and the like (167, 266). The

TABLE 11. Characterized species of spirochetes and relatives<sup>a</sup>

Spirochete subdivision	
<i>Spirochaeta</i>	<i>halophila</i>
<i>S.</i>	<i>aurantia</i>
<i>S.</i>	<i>litoralis</i>
<i>S.</i>	<i>isovalerica</i>
<i>Treponema</i>	<i>succinifaciens</i>
<i>T.</i>	<i>bryantii</i>
<i>T.</i>	<i>denticola</i>
<i>T.</i>	<i>phagadensis</i>
<i>S.</i>	<i>stenostrepta</i>
<i>S.</i>	<i>zuelzeri</i>
<i>T.</i>	<i>pallidum</i>
<i>Borrelia</i>	<i>hermsii</i>
<i>T.</i>	<i>hyodysenteriae</i>
Leptospira subdivision	
<i>Leptospira</i>	<i>patoc</i>
<i>Leptospira</i>	<i>interrogans</i>
<i>Leptothema</i>	<i>illini</i>

<sup>a</sup> Reference 167; Oyaizu, unpublished data; Weisburg et al., unpublished data. Indentation indicates specific relationship.

TABLE 12. Characterized species in the bacteroides-flavobacterium phylum (166, 234)\*

Bacteroides subdivision
<i>Bacteroides fragilis</i>
<i>B. ovatus</i>
<i>B. uniformis</i>
<i>B. asaccharolyticus</i>
<i>B. vulgatus</i>
<i>B. ruminicola</i>
<i>B. melaninogenicus</i>
<i>B. distasonis</i>
Flavobacterium subdivision
<i>Flavobacterium aquatile</i>
<i>Cytophaga johnsonae</i>
<i>Sporocytophaga myxococcoides</i>
<i>Cytophaga lytica</i>
<i>F. uliginosum</i>
<i>F. breve</i>
<i>Flavobacterium heparinum</i>
<i>F. ferrugineum</i>
<i>F. elegans</i>
<i>Saprospira grandis</i>
<i>Haliscomenobacter hydrossis</i>
Unnamed subdivision
Unnamed anaerobic flexible rod; strain PI-12fs

\* Indentation indicates specific relationship.

classical taxonomic distinction between spirochetes and treponemes, however, does not hold up (167); the two types form a genealogically intermixed unit. The lone species *Treponema hyodysenteriae*, however, represents a lineage distinct from the main spirochete-treponeme cluster (167), and the genus *Borrelia* is slightly peripheral to the main cluster as well (167).

It was previously suggested that the *Haloanaerobiaceae*, unusual anaerobic halophilic eubacteria, belong to this phylum (161), but full sequence information fails to confirm this (A. Oren and C. R. Woese, unpublished).

#### Bacteroides, Flavobacteria, and Their Relatives

The phylum made up of bacteroides and flavobacteria is an unexpected mixture of anaerobes, the bacteroides, and various aerobes, from genera such as *Flavobacterium*, *Cytophaga*, and others (166, 234). Table 12 lists some of its characterized representatives.

The grouping is cleanly defined by both sequence distance and parsimony analysis of rRNA sequences (234). Table 3 shows its quite distinctive signature. Note, for example, the U residue at position 570 (found in all members of this and the planctomyces phylum, but nowhere else among the eubacteria), the A residue at position 995 (otherwise a C, except in the green sulfur bacteria), and the A at position 1532 (which sets this group apart not only from all other eubacteria, but from archaeobacteria and eucaryotes as well [5, 166, 259]). Positions 570 and 866 are involved in a recognized "tertiary structural" interaction and so vary in concert (74).

A higher-order structural feature so far unique to this group is a series of three contiguous G-A pairs involving positions 1424-6 with 1474-6 (234). (Recall the two contiguous A-G pairs in this region characteristic of the gram-positive phylum; Fig. 10.) Common oligonucleotide sequences indicate that most if not all organisms in this phylum share the property (166).

The phylum's two major subdivisions separate the anaerobic *Bacteroides* species from the aerobic ones (166). The lack of phenotypic resemblance between members of the two subdivisions is remarkable, but may reflect only the fact that the bacteroides have been studied one way and the flavobacteria and relatives another. Most *Bacteroides* and (at least) a few species of *Flavobacterium* possess sphingolipids, compounds otherwise rare among eubacteria (234). An organism (yet unnamed) whose phenotype seems intermediate between these two has been isolated by K. O. Stetter (166); it is a strictly anaerobic flexible rod. Its phylogenetic position is also "intermediate" (166), and so the organism probably represents a third uncharacterized subdivision.

A sequence signature (derived from oligonucleotide catalogs) distinguishing the two subdivisions of the phylum is shown in Table 13. In most if not all cases the flavobacterial version would appear to be the ancestral one for the group, in that it is the one found in most of the other eubacterial phyla (166).

#### Planctomyces and Their Relatives

Species variously placed in the genera *Planctomyces*, *Pasteuria*, and *Pirella* (186, 202), together with the recently

TABLE 13. Sequence signature distinguishing the bacteroides-flavobacterium subdivisions\*

Position	Subdivision	Composition (%)				Consensus
		A	G	C	U	
306	B	0	100	0	0	A
	F	92	0	0	0	
316	B	0	0	0	100	C
	F	0	0	92	0	
337	B	100	0	0	0	G
	F	0	100	0	0	
718	B	0	0	9	91	A
	F	92	0	8	0	
809	B	55	9	0	0	G
	F	0	100	0	0	
947	B	100	0	0	0	G
	F	0	100	0	0	
986	B	0	0	100	0	A or U
	F	50	8	8	33	
1089	B	0	0	100	0	G
	F	?	>8	0	0	
1202	B	0	100	0	0	U
	F	0	0	0	100	
1234	B	0	0	0	100	C
	F	0	0	100	0	
1321	B	0	0	100	0	U
	F	0	0	0	92	
1356	B	0	0	100	0	G
	F	0	>8	0	16?	

\* References 166 and 234. See footnotes b to d in Table 10. B, Bacteroides subdivision; F, flavobacterium subdivision.

described hot spring organism *Isocystis pallida* (S. J. Giovannoni, in J. G. Holt, ed., *Bergey's Manual of Systematic Bacteriology*, vol. 3, in press; S. J. Giovannoni and H. Oyaizu, unpublished data), define this phylum. All are noted for the fact that their cell walls contain no peptidoglycan (110).

In terms of oligonucleotide catalogs, this is the most unique of eubacterial groups. Their  $S_{AB}$  values with other eubacteria are generally in the range of 0.10 to 0.15 (202), far lower than typical  $S_{AB}$  values between eubacterial phyla, which are normally 0.20 to 0.25 (Fig. 2). Oligonucleotide catalogs for the species in this group contain fewer of the highly conserved (ancestral) oligonucleotides than do those from any other phylum (202). As might then be expected, the group possesses a strong signature (Table 3).

The remarkable distance of the planctomyces group from other eubacteria measured by the rRNA cataloging approach was initially interpreted to mean that these organisms represent by far the deepest branching in the eubacterial line of descent (202). Analysis of full sequences, however, does not bear this out (H. Oyaizu and C. R. Woese, unpublished data); these large sequence distances are due to rapid evolution of the lineage, not an especially early divergence from the common line of eubacterial descent; see Fig. 11.

### Chlamydiae

The two known 16S rRNA sequences representing *Chlamydia*, i.e., from *Chlamydia psittaci* and *C. trachomatis*, are closely related; they show <5% difference (233). Since no other even moderately close relatives of these organisms are known, the phylum cannot yet be considered adequately described. A distant relationship between the chlamydiae and the planctomyces group is suggested by sequence signature (233). Of the sequence positions in Table 3 characteristic of the planctomyces and their relatives, five, i.e., 47, 48, 52, 53, and 353, are also found in the *C. psittaci* and *C. trachomatis* sequences (233). (Other positions in the 16S rRNA sequence, not shown in Table 3, that link the chlamydiae to the planctomyces and relatives are 110, 331, and 361 [233].)

Nevertheless, the chlamydiae should be considered to represent a distinct phylum, for the similarity between their 16S rRNAs and those of the planctomyces is too slight to place the two in the same taxon. Their sequence similarity, 71 to 72%, is considerably less than the 78% between *Planctomyces staley* and its relative *Isocystis pallida* (unpublished calculation). (The 71 to 72% figure is not the artificial result of a relatively rapid evolution in the chlamydial lineage, for chlamydial sequences are closer to outgroup sequences than are their counterparts in the planctomyces phylum. In addition, their  $S_{AB}$  values with other eubacteria are not as abnormally low as those of the planctomyces [202; unpublished calculation].) Interestingly, the chlamydiae, like the planctomyces group, also have cell walls that lack peptidoglycan (7, 59, 110).

### Radiation-Resistant Micrococci and Their Relatives

Until recently, radiation-resistant micrococci and their relatives was known to include only a few closely related species of radioresistant bacteria, i.e., *Deinococcus radiodurans* and its relatives (18, 56). However, it has now been shown to include the ubiquitous hot spring organism *Thermus aquaticus* as well (82). The signature shown for the phylum in Table 3 is rather weak. However, the two se-

quences now available (Weisburg, Ph.D. thesis; Giovannoni, unpublished data) suggest that the group should have a strong signature once this can be derived from full sequences rather than oligonucleotides. Similarity between the 16S rRNA sequences of *D. radiodurans* and *T. aquaticus* is 81% (unpublished analysis), low enough to place them in separate subdivisions of the same phylum.

### Green Non-Sulfur Bacteria and Their Relatives

The phylum containing the green non-sulfur bacteria is one of those for which little phenotypic justification exists. The group contains four known members, the thermophilic phototroph *Chloroflexus aurantiacus*, two mesophilic gliding species from the genus *Herpetosiphon*, and the thermophile *Thermomicrobium roseum* (63, 94, 162, 170). *Chloroflexus* and the green sulfur bacteria resemble one another in chlorosome structure and light-harvesting pigment type (63, 136); yet their rRNAs are unrelated (as mentioned above), and the structure of their photoreaction centers differs (45, 171). The unusual long-chain diols found in *Thermomicrobium*, functionally the equivalent of normal glycerol lipids (172), have recently been detected in *Chloroflexus* as well (T. A. Langworthy, personal communication), suggesting that a convincing phenotypic rationale for the grouping will ultimately be found.

Table 3 shows the group to have a fairly distinctive signature. The phylum is also characterized by higher-order 16S rRNA structural idiosyncracies (162). For example, the helical element between positions 1126 and 1146, a structure found in all other eubacterial sequences, is absent in the members of this group (162).

Although too few species have been characterized to project meaningful subdivisions, it would seem that *Thermomicrobium* represents one such and *Chloroflexus* and the *Herpetosiphon* species represent another (162). (Sequence homology between *Thermomicrobium* and the other species is a relatively low 77%.)

### Other Eubacterial Phyla

The 10 phyla described above account for almost all of the eubacterial species whose 16S rRNA have been cataloged or sequenced. Since the characterized strains are broadly representative of the known eubacteria, it might seem that few additional eubacterial phyla, if any, will be encountered. However, isolated rRNA sequences from several unusual eubacteria suggest that such is not the case, that many eubacterial phyla remain to be described, representing species yet to be discovered.

The two small clouds on this horizon, the two eubacteria whose 16S rRNA sequences do not belong to any of the above 10 phyla, are thermophiles noted for their unusual lipids. One is *Thermodesulfotobacterium commune* (unpublished catalog), a eubacterium having ether-linked lipids (119), while the other, *Thermotoga maritima*, has unique lipids that have so far defied complete characterization (89). The 16S rRNA sequence from *T. commune*, which is incomplete, will not be treated in this review; that of *Thermotoga maritima* (1), however, will play a key role in the subsequent discussion.

### Overall Structure of the Eubacterial Tree

The definition of the eubacterial phyla brings us to the limit of the rRNA cataloging approach. Unfortunately, this is

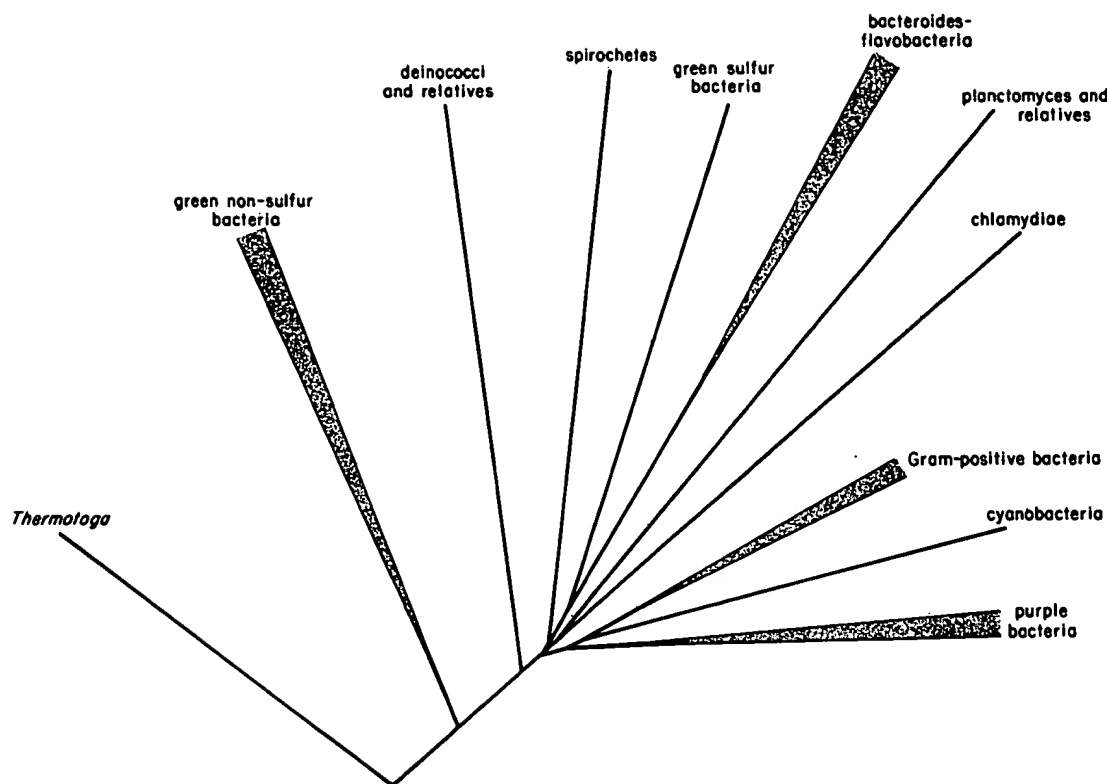


FIG. 11. Eubacterial phylogenetic tree based upon 16S rRNA sequence comparisons. An alignment was constructed (260) from one representative sequence from each of the eubacterial phyla together with an archaeobacterial consensus sequence (263). Using (only) those positions represented in all sequences in the alignment, a (corrected) evolutionary distance matrix was generated (99), from which a distance tree was constructed (36). Branch lengths on the tree are proportional to calculated distances. The sequences used are: *Thermotoga maritima* (1); green non-sulfur bacteria, *Thermomicrobium roseum* (162); deinococci and relatives, *Deinococcus radiodurans* (Weisburg, Ph.D. thesis); spirochetes, *Leptonema illini* (Oyaizu, unpublished data); green sulfur bacteria, *Chlorobium vibrioforme* (Weisburg, Ph.D. thesis); bacteroides-flavobacteria, *Flavobacterium heparinum* (234); planctomyces and relatives, *Planctomyces staley* (Oyaizu, unpublished data); chlamydiae, *Chlamydia psittaci* (233); gram-positive bacteria, *Bacillus subtilis* (68); cyanobacteria, *Anacystis nidulans* (224); purple bacteria, *Desulfovibrio desulfuricans* (163). For those phyla in which additional 16S rRNA sequences are available, the known sequence depth of the group has been (separately) calculated and is indicated by the shaded wedges. Additional sequences added to the alignment when calculating these depths are as follows: green non-sulfur bacteria, *Chloroflexus aurantiacus* (162); bacteroides-flavobacteria, *Bacteroides fragilis* (234); gram-positive bacteria, *Helicobacterium chlorum* (255); purple bacteria, *Agrobacterium tumefaciens* (274); and *Escherichia coli* (19).

the point at which the study of bacterial evolution starts to become interesting, for what we really want to know is the phylogenetic relationships among the phyla. With the aid of full 16S rRNA sequencing, it has become possible to resolve some of these branching orders, and so we can now begin to see the progression of eubacterial evolution.

Figure 11 shows the full eubacterial tree as it is presently known. This is a distance matrix tree (36), whose root has been subsequently determined by using an archaeobacterial consensus sequence as an outgroup (1). We shall discuss this tree and its implications in detail below, but two important points should be noted here. First, nine of the ten phyla described above, i.e., all except the green non-sulfur bacteria and relatives, appear to stem from roughly the same small region of the tree. Second, the present root of the tree separates all ten of the recognized phyla from the single species *Thermotoga maritima* (1). The microbiologist's at-

tention to date seems to have been confined to what phylogenetically is a nonrepresentative, sampling of the eubacteria: *Thermotoga* would appear to represent a vast unexplored "other world" of eubacteria, thermophilic organisms that (as their unique lipids suggest) almost certainly possess a variety of unusual and interesting biochemical properties and other characteristics.

It is important to stress that the root of the eubacterial tree shown in Fig. 11 is not an artifact of the treeing procedure. The same root placement results from distance-matrix and parsimony analyses. (Parsimony analysis on alignments of four sequences that comprise an outgroup sequence, the *Thermotoga* sequence, and all possible combinations of two other eubacterial sequences, show *T. maritima* in every case to cluster with the outgroup sequence, to the exclusion of the two other eubacteria [1].)

Sequence distance measurements (Fig. 5 and 11) show a

remarkable closeness of the *T. maritima* sequence to outgroup sequences, suggesting that its lineage is the most slowly evolving of all eubacterial lineages.

With the exception of the green non-sulfur group, all of the lineages branch from the eubacterial tree in such close proximity that their order of branching, the specific relationships among them, has yet to be convincingly determined. However, a few tentative relationships between various phyla are suggested and should be mentioned. A specific relationship may exist between the cyanobacteria and the gram-positive bacteria (Fig. 11). As we have seen, one signature position, 1207, in Table 3 supports this conjecture, for its composition is constant across all 150 or so gram-positive catalogs and all but one of the 30 or so cyanobacterial examples. That photosynthesis is found in both phyla and that chlorophyll *g* of *Heliobacterium* is most closely related in structure to chlorophyll *a* of the cyanobacteria (17) lend additional support to the possible relationship.

Another "superphylum" suggested in Fig. 11 involves the green sulfur bacteria and the bacteroides group. Table 3 shows several shared signature positions suggestive of that grouping, i.e., positions 995 and 1410, and a few higher-order structural features in 16S rRNA strengthen the case. The helix involving positions 1161 to 1175 (260) is altered in a way unique to these two phyla; one nucleotide is deleted from the loop in the vicinity of position 1167, while a "bulged" nucleotide is inserted in the stalk after position 1174; the structure in question can be seen in Fig. 7. (The deletion alone is seen in several other phyla, and the addition alone occurs in a particular subgroup of the  $\alpha$ -purple bacteria [234].) In both of these phyla the penultimate helix, positions 1435 to 1466, is strongly truncated, which is rare for eubacteria (unpublished analysis). The inclusion of the spirochetes, planctomyces, and chlamydiae in this superphylum is also suggested by the Fig. 11 analysis, but this relationship should not be considered seriously without additional evidence.

## ARCHAEBACTERIAL PHYLOGENY

### Unusual Nature of the Archaeobacterial Phenotype

Microbiologists have always perceived archaeobacteria as strange, highly atypical bacteria. Prior to their recognition as a phylogenetically coherent grouping (6, 261), however, their individual idiosyncrasies were interpreted merely as adaptations: the lipids of *Thermoplasma* were unusual because the organism evolved to live at high temperatures or in highly acidic environments or both (16); the wall of *Halococcus* was an adaptation to an extremely saline environment (123, 185); the uniqueness of their coenzymes merely reflected the capacity of methanogens to produce methane from carbon dioxide (277). That different archaeobacteria had the same unusual lipids was even interpreted as convergent adaptation (16)! As we find out more about the archaeobacteria our sense of their strangeness increases, but its explanation lies in a shared ancestry, not in individually evolved idiosyncrasies.

The archaeobacteria as we know them today are a collection of disparate phenotypes: the methanogens, the extreme halophiles, and the extremely thermophilic sulfur-metabolizing species (250, 270). Their metabolic differences are many, their known similarities few. The methanogens are noted for unusual coenzymes (Fig. 12), which tend not to occur in other bacteria. The extreme halophiles are the only photosynthetic archaeobacteria; they transduce light into chemical energy by means of a proton pump based on

bacteriorhodopsin, a pigment unique to this group of organisms (122, 217). The halobacteria also possess remarkably high intracellular salt concentrations (113). (Some methanogens have impressively high internal salt concentrations as well [95], but not in the range characteristic of the halophiles.) The extreme thermophiles also have at least one unique coenzyme (28, 34) (Fig. 12). They share with methanogens a capacity to reduce large amounts of sulfur (213). The extreme thermophiles have no immediate known relatives that grow at or near normal temperatures; most of them grow best at remarkably high temperatures (212, 215). Systematic comparisons of archaeobacterial metabolism are definitely needed.

The branched-chain, ether-linked lipids common to all archaeobacteria are found nowhere else in nature (35, 69, 103, 117-120). In many, but not all species the glycerol diethers tend to be covalently joined "head to head," to produce diglycerol tetraethers, which form unusual membranes that cannot be freeze-fractured (118). Archaeobacteria show an impressive number of variations in lipid structure based upon the ether-linked, branched-chain theme (35, 69, 117-119).

Archaeobacteria display their own characteristic version of every major macromolecular function, e.g., 16S rRNA (Fig. 6). However, within these versions an impressive spectrum of variation can occur (250). For example, unlike eubacteria, their walls are varied in type (101, 109). In their structural details some archaeobacterial 5S rRNAs resemble somewhat the eubacterial form of the molecule and others resemble the eucaryotic form, while still others are unique (53, 133, 208). All three kingdoms have a characteristic subunit pattern for DNA-dependent RNA polymerases, but within the archaeobacteria relatively drastic variations in type are seen (187, 282). For the extremely thermophilic species and *Thermoplasma acidophilum*, the so-called B subunit is a single large molecule, while for methanogens and their relatives it consists of two smaller molecules, B' and B'' (187, 282). In their transfer RNAs (tRNAs) archaeobacteria show a characteristic pattern of modified bases (70, 71, 73, 146, 165, 273), yet the tRNAs of the extreme thermophiles (except for *Thermoplasma*) are much more highly modified than are those of the methanogens and relatives (250, 259).

No new major archaeobacterial phenotypes were discovered in the 9 years following the recognition of archaeobacteria as a distinct group, which leads to a growing feeling that the three basic and highly unique phenotypes, methanogenesis, extreme halophilism, and extremely thermophilic sulfur metabolism, are all that exist in the kingdom. (However, see below.)

### Definition of the Major Archaeobacterial Groups

The number of archaeobacterial species characterized by rRNA analysis is only one-tenth the number of characterized eubacterial species. Although their number is not large enough to provide phylogenetic detail, it is sufficient to identify the higher archaeobacterial taxa, for the sampling is broadly representative.

rRNA cataloging studies showed that there are three major groups of methanogens and one of extreme halophiles (5, 52, 54, 56, 250), each being the equivalent of a eubacterial phylum. Although the branching order among the four phyla could not be determined by this method, the four as a group were easily distinguished from the group of extremely thermophilic archaeobacteria (excluding *Thermoplasma*). However, too few species of extreme thermophiles were

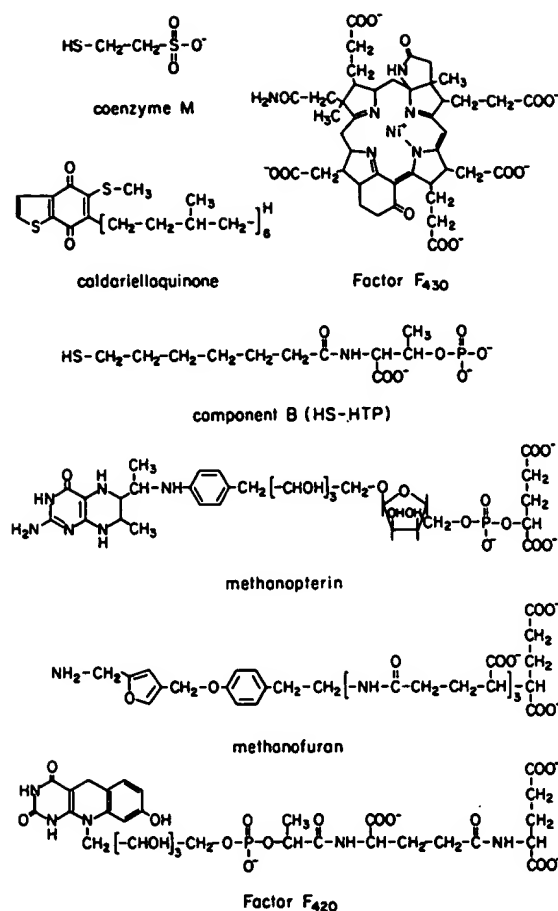


FIG. 12. Structure of various archaeobacterial cofactors (34, 154, 238). All are found in methanogens and utilized in methane production, except for caldariella quinone, which occurs in *Sulfolobus* species.

characterized by the cataloging method to give a sense of that group's phylogeny (225, 259).

Table 14 lists the archaeobacterial species that have been characterized by the rRNA method, in an approximate phylogenetic ordering. Tables 15 and 16 define the major archaeobacterial groupings by sequence signature.

**Methanogens.** The three methanogen phyla defined by rRNA sequence comparisons can also be recognized by morphological criteria, with a few exceptions (5, 238). Immunological cross-reactivities almost always identify methanogen group affiliation as well (135). As a result, we know that each of the three main groups of methanogens (currently designated orders), *Methanobacteriales*, *Methanococcales*, and *Methanomicrobiales*, contains a large number of species (5, 98, 135, 238, 239). To date, the immunological studies have given no clear indication that any additional major methanogen group exists (135). The three methanogen phyla are also distinguishable by 5S rRNA type (53).

In two of the methanogen phyla, subdivisions are recognized by rRNA cataloging. The *Methanobacteriales* breaks naturally into two "genera," *Methanobacterium* and *Meth-*

*anobrevibacter* (5, 238). A third genus, now represented by the lone species *Methanothermobacter fervidus* (214), should ultimately be declared a distinct unit within this phylum, or perhaps a separate phylum (238). The *Methanomicrobiales* in turn divide into two distinct subdivisions, (formally families) the *Methanomicrobiaceae* and the *Methanosarcinaceae* (5, 238). The latter is the most metabolically unusual of the methanogen groups. Its species can utilize acetate or sometimes methyl amines in methane production; some are even unable to use carbon dioxide for this purpose (9, 42, 153). These are also the only methanogen species that contain cytochromes (b or c or both) (112). The unusual halophilic methanogens belong to this group as well (13; I. M. Mathrani, D. R. Boone, and R. A. Mah, Abstr. Annu. Meet. Am. Soc. Microbiol. 1985, 185, p. 160).

**Extreme halophiles.** The extreme halophiles constitute one

TABLE 14. Archaeobacterial subdivisions, representative genera, and species<sup>a</sup>

<b>Methanococcus group</b>
<i>Methanococcus vannielii</i>
<i>M. voltae</i>
<i>M. maripaludis</i>
<i>M. thermolithotrophicus</i>
<i>M. jannaschii</i>
<b>Methanobacter group</b>
<i>Methanobacterium formicicum</i>
<i>M. bryantii</i>
<i>M. thermoautotrophicum</i>
<i>Methanobrevibacter smithii</i>
<i>M. arborophilus</i>
<i>M. ruminantium</i>
<i>Methanosphaera stadtmaniae</i>
<i>Methanothermobacter fervidus</i>
<b>Methanomicrobium group</b>
<i>Methanosarcina barkeri</i>
<i>Methanococcoides methylutens</i>
<i>Methanothermobacter soehngenii</i>
<i>Methanospirillum hungatei</i>
<i>Methanomicrobium mobile</i>
<i>Methanomicrobium paynteri</i>
<i>Methanogenium cariaci</i>
<i>Methanogenium marisnigri</i>
<i>Methanoplanus limicola</i>
<b>Halobacteria</b>
<i>Halobacterium volcanii</i>
<i>H. cutirubrum</i> <sup>b</sup>
<i>H. sodomense</i>
<i>H. trypanicum</i>
<i>Halococcus morrhuae</i>
<b>Thermoplasma</b>
<i>Thermoplasma acidophilum</i>
<b>Thermococcus group</b>
<i>Thermococcus celer</i>
<b>Extreme thermophiles</b>
<i>Sulfolobus solfataricus</i>
<i>S. acidocaldarius</i>
<i>Thermoproteus tenax</i>
<i>Desulfurococcus mobilis</i>
<i>Pyrodicticum occultum</i>

<sup>a</sup> References 5, 98, 238, and 263. Approximate phylogenetic clustering suggested by indentation.

<sup>b</sup> 16S rRNA identical to those of *H. salinarum* and *H. halobium*.



of the most distinctive groups of bacteria known. As mentioned, both their very high internal salt concentrations (i.e., in the range of 5 M potassium ion [113]) and their mechanism for photoproduction of energy are unique in the bacterial world. (Bacteriorhodopsin does superficially resemble the eucaryotic visual pigment, however, hence its name [122, 217].) Over ten species of extreme halophiles have been characterized in terms of 16S rRNAs, and they form a relatively close-knit grouping (56, 72, 90, 125; C. R. Woese and G. E. Fox, unpublished data). The unusual halophiles that grow under alkaline conditions (222) are among them. The group is not a particularly deep one. Based upon relative  $S_{AB}$  values (Woese and Fox, unpublished data) and known sequences, all rRNAs in the group show at least 87% sequence similarity. The internal phylogenetic structure of the group is unspectacular.

The extreme halophiles are known to contain cytochromes and ferredoxins (78, 113, 121). It has been reported that halophile ferredoxin sequences are specifically related to those found in cyanobacteria, to the exclusion of other eubacteria (80). If such homology exists, it is highly unlikely to reflect a genuine phylogenetic relationship; the cyanobacteria and extreme halophiles are not related to one another to the exclusion of the other members of their respective kingdoms. Sequence convergence also seems unlikely. However, gene transfer does not, for the two types of organisms can in some cases share the same habitat.

The cataloging approach failed to distinguish the branching order among the three groups of methanogens and the extreme halophiles. Most microbiologists tacitly assumed that the four phyla were arranged along phenotypic lines; i.e., all methanogens clustered together to the exclusion of the halophiles. However, oligonucleotide signatures weakly suggested a specific relationship between the halophiles and the *Methanomicrobiales* (275).

**Extreme thermophiles.** The extreme thermophiles are the least characterized, but (as will be seen below) the most evolutionarily interesting, of the archaeobacteria. The extreme thermophiles seem quite uniform in phenotype. All species grow anaerobically, and most require sulfur as an energy source (215). A minority of species can also grow aerobically, and some that use sulfur as an energy source do not require it (215). Most species grow best at extremely high temperatures, some near the boiling point of water (212).

The extreme thermophiles differ from other archaeobacteria in numerous ways. Their modes of division tend to be unusual and varied (281), as do their viruses (279, 281); their ribosomal subunits have an unusual shape (see discussion below) (81); they have (as mentioned) at least one unique coenzyme; they seem to be insensitive to most antibiotics (22); and both their tRNAs and rRNAs are highly modified (the level in the latter case is fivefold greater than that seen in the methanogens and their relatives) (259). Nevertheless, in sequence terms, in membrane structure, and in most phenotypic characteristics, the extreme thermophiles are definitely archaeobacteria (141, 142, 215, 263).

The phenotypic clustering of extreme thermophiles is deceptive. As we shall see, it is not a phylogenetic clustering. rRNA cataloging studies showed that *Thermoplasma* was more closely related to the methanogens and relatives than to the extreme thermophiles (56). Even its pattern of 16S rRNA base modification is "methanogen-like" (56, 262). However, this potential relationship was at first given little credence, and the organism tended to be viewed as an atypical extreme thermophile (215). Although an rRNA

catalog did not exist for *Thermococcus*, there seemed every reason to assume it was related to the other extreme thermophiles (215), but, as we shall see, this is not so.

### The Archaeobacterial Tree: Its Branching Order and Root

An evolutionary distance matrix for the known archaeobacterial 16S rRNA sequences is shown in Table 17. A phylogenetic tree derived from evolutionary distances is shown in Fig. 13 (263). With two minor exceptions, the same branching order is given by a number of different methods: parsimony analysis, distance-matrix treeing, using subsets of the positions in the sequence alignment, etc. (263; unpublished analysis). One of the exceptions, *Sulfolobus solfataricus*, tends to cluster with *Desulfurococcus mobilis* by parsimony analysis and when no or few outgroups are included in a distance treeing analysis; yet when the full set of archaeobacterial sequences is analyzed by a distance-matrix method, *D. mobilis* clusters instead with *Pyrodicticum occultum*, to the exclusion of *S. solfataricus*. The *S. solfataricus* line of descent is more rapidly evolving than are these others (as can be inferred from Table 17), which tends to force its branching deeper into the tree than it actually is, particularly when outgroup sequences are included in the analysis.

For the same reason the exact branching order for *Thermoplasma acidophilum*, the second exception, is uncertain, though its lineage always remains in the general vicinity of the position shown in Fig. 13.

**Main branches.** The root of the archaeobacterial tree divides the urkingdom into two main lineages: a cluster of extreme thermophiles on the one hand, and the methanogens and their relatives on the other.

The two main branches differ from each other in several interesting ways. For one, the methanogen branch appears to be the phylogenetically "deeper" of the two. For another, the extreme thermophile branch (so far) is phenotypically pure, whereas its counterpart is cosmopolitan, including a mixture of methanogenic, extremely halophilic, and extremely thermophilic phenotypes.

TABLE 15. Sequence signature distinguishing the two main archaeobacterial branches

Position <sup>a</sup>	Composition in:			
	Methanogens and relatives <sup>b</sup>	<i>Thermoplasma acidophilum</i> <sup>c</sup>	<i>Thermococcus celer</i> <sup>d</sup>	Sulfur-dependent archaeobacteria <sup>e</sup>
34	U	?	—	C
242	G	U	—	C
559	U <sup>f</sup>	—	—	G
939	C	—	G	G
965	Y	—	—	G
1074	A	—	—	G
1088	U	A	A	G
1252	U	—	—	C
1351	U	—	C	C
1408	A	—	—	G

<sup>a</sup> In 16S rRNA sequence (260).

<sup>b</sup> Based upon oligonucleotide catalogs or sequences for 21 species of methanogens, and 9 species of halophiles (5, 6, 54, 72, 91, 125; Woese, unpublished data).

<sup>c</sup> Based upon catalog (262) and an unpublished sequence (R. A. Zimmermann, personal communication).

<sup>d</sup> Based upon the sequence (Woese et al., unpublished data).

<sup>e</sup> Based upon oligonucleotide catalogs or sequences for five sulfur-dependent archaeobacteria (126, 158, 225; R. Garrett, personal communication; Woese et al., unpublished data).

<sup>f</sup> —, Same composition as in methanogens and relatives.

<sup>g</sup> One example of G at this position (5).

TABLE 16. Sequence signatures defining the methanogen and extreme halophile groups<sup>a</sup>

Group	Position <sup>b</sup>	Major base	Composition in:					
			Mco <sup>c</sup>	Mba <sup>d</sup>	Mmi <sup>e</sup>	Hal <sup>f</sup>	Tac <sup>g</sup>	Tce <sup>h</sup>
<i>Methanococcus</i>	319	A	G	— <sup>i</sup>	—	—	—	—
	334	U	C	—	—	—	—	A
	703	A	G	—	—	—	—	—
	976	A	G	—	—	—	—	G
	1122	Y	A	—	—	—	—	—
	1150	U	A	—	—	—	—	—
	1366	G	C	—	—	—	C	—
	1377	A	G	—	—	—	—	—
<i>Methanobacterium</i> and relatives	1127	G	—	A	—	—	—	—
	1167	C	—	A	Y	—	—	A
	1342	C	—	G	—	—	—	—
	1357	U	—	G	—	—	G	G
	1365	A	—	C	—	—	C	C
	1417	G	—	A	—	—	—	—
<i>Methanospirillum</i> and relatives	271	C	—	—	A	—	—	—
	366	C	—	—	A	—	—	—
	807	U	—	—	C	—	—	—
	1076	C	—	—	U	—	—	—
	1084	G	—	—	A	—	—	—
	1122	U	A	—	C	—	—	C
Extreme halophiles	33	C	—	—	—	U	?	—
	50	A	—	—	—	Y	?	—
	274	A	—	—	—	U	—	—
	374	A	—	—	—	C	—	—
	940	G	—	—	—	C	—	—
	1291	C	—	—	—	U	—	—
	1317	U	—	—	—	A	—	—
	1343	C	—	—	—	G	—	—

<sup>a</sup> See footnotes to Table 15 for details.<sup>b</sup> In 16S rRNA sequence.<sup>c</sup> Mco, *Methanococcus*, five species (5, 96, 98).<sup>d</sup> Mba, *Methanobacter* and relatives, eight species (5, 124; unpublished data).<sup>e</sup> Mmi, *Methanospirillum* and relatives, eight species (5, 275; unpublished data).<sup>f</sup> Hal, Extreme halophiles, nine species (72, 91, 125; unpublished data).<sup>g</sup> Tac, *Thermoplasma acidophilum* (K. M. Cao, H. Ree, D. L. Thurlow, and R. A. Zimmermann, personal communication).<sup>h</sup> Tce, *Thermococcus celer* (unpublished data).<sup>i</sup> —, Composition equals major base.TABLE 17. Percent similarities and evolutionary distances for archaeobacterial 16S rRNAs<sup>a</sup>

Species	<i>Mc. vannielii</i>	<i>M. formicicum</i>	<i>Ms. hungatei</i>	<i>Methanosarcina</i> sp. strain WH-1	<i>H. volcanii</i>	<i>T. acidophilum</i>	<i>Tc. celer</i>	<i>S. solfataricus</i>	<i>Tp. tenax</i>	<i>D. mobilis</i>	<i>P. occultum</i>
<i>Mc. vannielii</i>		20.3 <sup>b</sup>	27.7	26.4	27.8	30.0	19.7	28.6	29.0	24.2	25.5
<i>M. formicicum</i>	69.2 <sup>c</sup>		23.0	25.0	25.5	29.3	20.0	30.0	28.0	24.1	24.8
<i>Ms. hungatei</i>	59.9	65.7		19.8	23.6	33.6	26.3	33.6	32.9	30.4	30.8
<i>Mr. sp. strain WH-1</i>	61.5	63.2	69.8		26.3	30.0	24.7	32.3	30.9	29.8	28.5
<i>H. volcanii</i>	59.7	62.6	64.9	61.6		31.5	27.2	35.6	33.9	32.2	30.8
<i>T. acidophilum</i>	57.2	58.0	53.1	57.2	55.5		26.4	33.9	34.6	33.3	31.9
<i>Tc. celer</i>	70.0	69.5	61.6	63.5	60.5	61.5		24.5	20.0	19.1	17.7
<i>S. solfataricus</i>	58.8	57.2	53.1	54.5	50.9	52.8	63.8		17.1	12.4	12.8
<i>Tp. tenax</i>	58.3	59.6	53.9	56.1	52.8	52.0	69.5	73.5		14.1	11.7
<i>D. mobilis</i>	64.1	64.3	56.7	57.3	54.7	53.4	70.8	80.3	77.7		6.9
<i>P. occultum</i>	62.6	63.3	56.2	58.9	56.2	55.0	72.7	79.6	81.2	88.6	

<sup>a</sup> Sequence references are cited in Fig. 13 legend. *Mc.*, *Methanococcus*; *M.*, *Methanobacterium*; *Ms.*, *Methanospirillum*; *Mr.*, *Methanosarcina*; *H.*, *Halobacterium*; *T.*, *Thermoplasma*; *Tc.*, *Thermococcus*; *S.*, *Sulfolobus*; *Tp.*, *Thermoproteus*; *D.*, *Desulfurococcus*; *P.*, *Pyrodictum*.<sup>b</sup> Upper right numbers are evolutionary distances (99); only positions in alignment represented in all sequences are used in calculation.<sup>c</sup> Lower left numbers are percent similarity; all positions not represented in all sequences and all positions of constant composition were removed from consideration. (This last has been done to accentuate the differences among sequences; it does not change their rank order.)

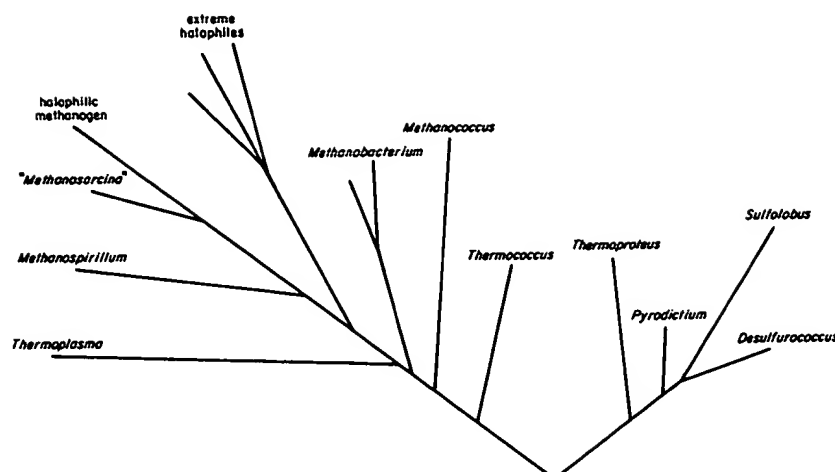


FIG. 13. Archaeobacterial phylogenetic tree based upon 16S rRNA sequence comparisons. The sequences listed were aligned and an unrooted distance tree was constructed as in the legend to Fig. 11. Its root was subsequently imposed on the basis of outgroup consensus sequences (eubacterial and eucaryotic); the root given by eubacteria or eucaryotes is in the same general region of the tree (i.e., between *Thermococcus* and the other extremely thermophilic species [263], and that shown represents an average of the eubacterial and eucaryotic placements. Those sequences used in the alignment are *Methanospirillum hungatei* (275); the halobacteria *Halobacterium volcanii* (72), *Halococcus morrhuae* (125), and *Halobacterium cutirubrum* (91), from left to right in that order; *Methanosarcina* sp. strain WH-1 (P. Rouviere, unpublished data; Woese et al., unpublished data); halophilic methanogen strain FS-1 (Mathrani et al., Abstr. Annu. Meet. Am. Soc. Microbiol. 1985; Woese et al., unpublished data), *Methanobacterium thermoautotrophicum* (R. Garrett, personal communication) and *Methanobacterium formicicum* (124), from left to right, representing *Methanobacterium*; *Methanococcus vannielii* (96); *Thermococcus celer* (Woese et al., unpublished data); *Thermoproteus tenax* (126); *Pyrodictium occultum* (Woese et al., unpublished data); *Sulfolobus solfataricus* (158); and *Desulfurococcus mobilis* (Garrett, personal communication).

***Thermococcus* and *Thermoplasma*.** The two extreme thermophiles *Thermococcus* and *Thermoplasma* are unrelated to their phenotypic counterparts. As we have seen, their phylogenetic placement with the methanogens is consistent in the case of *Thermoplasma acidophilum* with the pattern of modified nucleotides in the 16S rRNA, i.e., a low level of modifications, at particular sites (5, 262). (Nothing is known yet about the pattern of modified nucleotides in *Thermococcus* RNAs.)

Interestingly, *Thermococcus celer* is not closer to the methanogens than to the other extreme thermophiles by overall sequence distance measure. Table 17 shows its 16S rRNA to be closest to that of *Pyrodictium occultum*. This would seem, however, to reflect the relatively slow evolutionary tempo among extreme thermophiles in general, not a specific relationship between these two organisms, a point that can be clearly demonstrated by signature analysis, which focuses on the more conserved positions in the molecule. For example, in a 16S rRNA alignment containing sequences from eight methanogens, three extreme halophiles, and four representatives of the extreme thermophile branch, there are about 30 positions that have a constant composition among the methanogens and extreme halophiles but a different (constant) composition among the extreme thermophiles. The *Thermococcus celer* sequence exhibits the characteristic methanogen composition in about 79% of these cases; the extreme thermophile composition in only 11% (263; Woese, unpublished analysis).

That *Thermococcus* and *Thermoplasma* resemble other extremely thermophilic archaeobacteria in DNA-dependent RNA polymerase subunit pattern (187, 282) does not necessarily prove genealogical relatedness. Zillig and co-workers have shown that the large B subunit of RNA polymerase,

found in all extreme thermophiles, is undoubtedly the ancestral type, for the smaller B' and B'' subunits appear to have arisen from it at least twice (281). For this reason the common occurrence of a large RNA polymerase B subunit does not necessarily mean specific relationship.

The branching of *Thermococcus* from the main methanogen line of descent is sufficiently deep to suggest that it may ultimately be considered to represent a third major archaeobacterial lineage.

**Branching order among the methanogens and extreme halophiles.** Full sequences show that the extreme halophiles cluster specifically with the *Methanomicrobiales*, to the exclusion of the other two phyla of methanogens, a relationship that was hinted at by oligonucleotide signatures (263, 275). Although this is an unexpected, even counterintuitive, finding, the evidence supporting it is entirely convincing. The conclusion readily emerges from parsimony analyses of 16S rRNAs as well as distance treeing (263). For example, in an alignment that includes sequences from the three extreme halophiles, one from each of the methanogen phyla, and the *Thermococcus* sequence, there are 22 positions having a common composition in the halophile and *Methanospirillum* sequences that have a different common composition in the remaining three sequences. For any other combination of these sequences (keeping the three halophile sequences as a unit), there are no more than seven positions of common composition defined in this way (263; unpublished analysis).

Although the relationship between the extreme halophiles and the *Methanomicrobiales* is difficult to justify phenotypically, certain facts are consistent with such a grouping. For example, methanogens capable of growth at high salt concentrations belong to the *Methanomicrobiales* (Fig. 13). On the methanogen branch of the tree, it is only among species

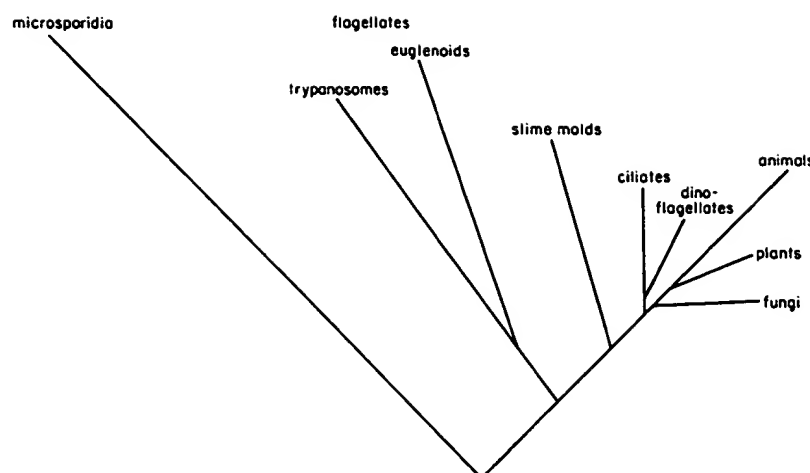


FIG. 14. Eucaryotic phylogenetic tree based upon 16S-like rRNA sequence comparisons. The eucaryotic sequences listed were aligned and a distance matrix tree was constructed as in the legend to Fig. 11. The root was determined by including an archaeobacterial consensus sequence in the alignment. Those eucaryotic sequences used are as follows: microsporidia, *Vairimorpha necatrix* (226a); flagellates, *Trypanosoma brucei* and *Euglena gracilis* (196); slime molds, *Dictyostelium discoideum* (145); ciliates, *Paramecium tetraurelia* (195); dinoflagellates, *Prorocentrum micans* (84); fungi, *Saccharomyces cerevisiae* (179); animals, *Xenopus laevis* (181); plants, *Zea mays* (147). The archaeobacterial (100%) consensus sequence was derived from an alignment of three methanogens and three extreme thermophiles.

of *Methanosarcina*, the extreme halophiles, and *Thermoplasma* that cytochrome *b* or *c* or both are found (78, 86, 112, 121, 190).

Figure 13 also indicates a specific relationship between the *Methanobacteriales* and the halophile-*Methanomicrobiales* group, to the exclusion of the *Methanococcales*. This relationship, too, can be rationalized by sequence signature (263). For example, the *Methanobacterium formicicum* sequence shares more than twice as many positions of exclusive common sequence with the *Methanospirillum hungatei* and extreme halophile sequences as does the *Methanococcus vannielii* sequence (263).

The position of *Thermoplasma* in the tree, as mentioned, is uncertain. Various treeing procedures place its branching in a range that extends from somewhere on the common *Methanospirillum*-halophile branch to just below the *Methanococcus* branch (263). A signature marginally suggestive of the first placement exists (263), but the exact genealogy of the organism should be considered uncertain.

#### A BRIEF LOOK AT EUCARYOTE PHYLOGENY

The biologist generally feels that he has a relatively good sense of eucaryotic evolution, and up to a point this is certainly true. Detailed taxonomies exist for the various classes of animals and plants that rather accurately reflect their phylogenies. The higher levels of eucaryote classification are another matter, however. Here our understanding is no less a matter of prejudice and preconception than it was for the bacteria. The so-called Five Kingdom classification (139, 240, 241), plants, animals, fungi, protists, and monera, cannot be considered proper phylogeny. It mixes apples and oranges and defines categories by exclusion. The system gives the same kingdom rank to each of the four groups of eucaryotes that it gives to one group of procaryotes (monera). Yet it has been obvious for some time that the four eucaryotic kingdoms form a phylogenetically coherent unit

that as a whole ranks with the monera; and monera, of course, is now known to comprise two separate kingdoms (56). Within the eucaryotes, the protists do not form a phylogenetically coherent unit (196).

Fortunately, this scheme and our criticism of it will soon be rendered academic. A new and very different view of eucaryotic phylogeny is beginning to emerge, whose outlines can be seen in Fig. 14.

#### Implications of the Eucaryote Phylogenetic Tree

What this preliminary phylogeny begins to suggest is that the major epochs in eucaryote evolution corresponded to major periods in earth history. A relatively "recent" period of massive evolutionary radiation appears to have given rise to most of the major eucaryotic lineages: green plants, animals, fungi, ciliates, dinoflagellates, (some) amoebae, etc. (196). (The cellular slime molds represent a slightly earlier branching [145, 196].) It is tempting to equate the onset of this particular radiation with a globally significant event and attribute the radiation to some major innovative biological (evolutionary) response. The hydrosphere is thought to have become oxidizing about 1.5 billion years ago (228), which time is roughly consistent with the occurrence of the radiation estimated by back-extrapolation from known time points on the eucaryotic tree (196). Biologically this could have been a time when an oxygen-utilizing mitochondrion developed (274).

Deeper lineages occur on the eucaryotic tree, however, that appear to predate significantly this period of massive radiation. A group of flagellates is seen to branch from the main eucaryotic stem well before the radiation in question (196), and the microsporidian lineage definitely emerges well before that (226a). These earlier branchings should reflect earlier phases in the planet's history. The flagellate branching might stem from the "microaerobic" or "amphibiaerobic" period, i.e., the era between the time the atmosphere be-

came significantly oxidizing (over 2 billion years ago) and the hydrosphere became aerobic (about 1.5 billion years ago) (228). The microsporidian branching would represent the still earlier genuinely anaerobic period, before 2.5 or so billion years ago (226a).

Microsporidia, e.g., the genus *Vairimorpha*, are a group of highly unusual and little studied unicellular eucaryotes, deserving of more than passing attention. They have primitive modes of cell division (175) and exhibit strange and interesting life cycles, connected with their obligately parasitic mode of existence. The group as a whole parasitizes an extremely wide range of other eucaryotes; they have been seen to infect examples of all the animal phyla, and they even parasitize other protists (197). This could be interpreted to mean that their parasitism is of very ancient origin. Microsporidia have no mitochondria and, given their deep phylogenetic branching, might never have had them.

Perhaps their molecular idiosyncrasies are the most fascinating aspect of the microsporidia (and very little is known about this). These are the only eucaryotes that have no 5.8S rRNA (227). Their rRNAs are also far smaller than normal eucaryotic rRNAs. Typically, the eucaryotic small-subunit rRNA comprises about 1,800 nucleotides, but its microsporidian counterpart contains under 1,300 nucleotides, even less than the roughly 1,450 to 1,550 nucleotides characteristic of procaryotes (75, 226a). Interestingly, the missing areas in the microsporidian small-subunit rRNA tend to be those that are unique to the eucaryotes (75) (Fig. 5).

With the microsporidian branching we may be near the base of the eucaryotic tree, the beginnings of eucaryotic cellular evolution. Microsporidia are defined as eucaryotes because they have a membrane-delimited nucleus. The question is, in what other ways do they resemble eucaryotes, and what characteristic eucaryotic features do they lack? Eucaryotes would appear to be an old group, far older than many biologists might have thought. Their antiquity would seem to rival that of the procaryotic kingdoms.

## NATURE OF THE EVOLUTIONARY PROCESS IN BACTERIA

### Relationship Between Evolution's Tempo and Its Mode

Classical evolutionists recognized that a relationship existed between the rate at which evolution proceeds, its tempo, and the quality of the changes that occurred, its mode (143, 192). Fossil evidence showed that some lineages evolve more rapidly than others and that rates of phenotypic change vary within lineages at different stages in their history (192). Evolution tended to be particularly rapid as a lineage came into being and also in some cases as it died out (192). The quality of phenotypic change was different during such periods of rapid evolution; it was often described as drastic and novel, even bizarre (192).

Two general rules governed the relationship between the tempo and mode of metazoan evolution: (i) true evolutionary novelty (of the kind that gives rise to major groups) occurred only during times of rapid evolution, and (ii) rapid evolution tended to be episodic, not chronic. (Evolution of the horse was often used as an example. A relatively short period of dramatic evolutionary change transformed the common ancestor of horses, tapirs, and rhinoceroses into a horselike creature. The evolution of this ancestral horse into the modern form, which was a far more protracted affair,

involved relatively little further change [192].) Two other characteristics of rapid evolution are its instability and radiation. Of the numerous lineages typically formed from a common ancestral stock during these saltatory episodes, many, if not most, were short-lived (192). The origin of the animal phyla conformed to such a pattern—all seem to have burst forth, almost simultaneously (192). The same can now be said of the origin of the main eucaryotic kingdoms (Fig. 14).

Evolutionists have debated the whys and wherefores of the tempo-mode issue for decades. In the past discussion centered about whether or not the same evolutionary mechanisms or environmental conditions underlay the chronic progressive evolution which characterized normal established lineages, usually referred to as microevolution, that underlay rapid episodic evolution, variously called macroevolution, megaevolution, or quantum evolution (192). Initially, global catastrophes or elevated mutation rates had been invoked to explain the radiating, saltatory origin of major taxa, and some biologists went so far as to declare macroevolution (megaevolution) and microevolution to be different in kind (66). However, the idea that catastrophes played a necessary or even major role in radical evolutionary change was later rejected (192), and elevated mutation rates were no longer seen as required for episodic, drastic evolutionary change; in fact, they could not explain it (192). Rapid evolution, macroevolution, was solely the result of ecological considerations, population genetics: small population sizes, rapidly changing environments, untoward conditions, and the like, were all that need be invoked (143, 192, 272). (However, note that we are seeing a return in recent times to the global catastrophe type of explanation for radical evolutionary change; good evidence now supports the idea that the effects of a comet or asteroid impact lead to the extinction of the dinosaurs, thereby allowing the subsequent evolutionary radiation of the mammals [3].)

Some biologists today seem to feel that microevolution-macroevolution is a nonissue, the difference between them being only a matter of degree. All distinctions, all boundaries, however, are matters of degree when viewed finely enough; this is especially apparent for protracted processes such as evolution. The significance of a distinction turns not on whether it is a matter of degree, but on how sharp the boundary is relative to the space/time scale of the phenomena that define it.

One thing is clear about the tempo-mode issue: it will never be resolved if its study is confined to fossil evidence. In these terms the crucial parameters are too poorly defined and distinguished; the phenomena are too illusive and inaccessible and too difficult to explain. It is important, therefore, to try to generalize the tempo-mode problem to bacterial systems and to study it in molecular terms.

(In the following discussion I will use the term "macroevolution" to mean the episodic, saltatory, radiating type of evolution that can create major taxa and is often associated with instability in the newly formed lineages [192]. Although this may not be the strict conventional usage of the term, there should be no problem as long as its usage is understood.)

### Macroevolution at the Molecular Level: the Mycoplasmas and Their rRNAs

Changes in molecular sequence are the most basic manifestation of evolution's tempo. Molecular chronometers, as



FIG. 15. Phylogenetic tree for the mycoplasmas and other members of the gram-positive bacteria (see Table 18 footnotes), based upon 16S rRNA sequence comparisons. An alignment consisting of the sequences shown was used to construct an unrooted distance matrix tree as in the legend to Fig. 11. The root was imposed by using these sequences: *Anacystis nidulans* (224), *Desulfovibrio desulfuricans* (163), *Agrobacterium tumefaciens* (274), and *Escherichia coli* (19).

we have seen, can measure a pure tempo, essentially unaffected by the mode, the overlying phenotypic changes. Unfortunately, sequence distances do not provide as good a measure of evolutionary rate as we would like. They provide only average rates, over relatively long evolutionary time spans, when what is required for a proper formulation of the tempo-mode problem in molecular terms are measures of changes in rate during the course of a lineage's evolution, as are found in the metazoan fossil record. However, as will be seen, the "chronometric structure" of rRNAs (and presumably other macromolecules) is such that not only average rates, but also indications of rate changes (i.e., peak rates) are recorded. The main problem we appear to face, then, is finding the molecular counterpart of that ill-defined quality, evolutionary mode, if such exists. This problem too will prove tractable, for on the molecular level evolutionary tempo and mode are intimately connected; they are essentially different manifestations of the same process.

If macroevolution occurs among bacteria, it should be most evident in the most rapidly evolving bacterial lineages. Two eubacterial lineages in particular are attractive from this point of view: the planctomyces group (202) and the mycoplasmas. Of these, the mycoplasmas are at least as rapidly evolving as the planctomyces. And, what is more important for the present discussion, they have known close relatives that evolve at normal rates, e.g., *Bacillus* (262). Our discussion will therefore focus on mycoplasmas.

Mycoplasmas show the main characteristic expected of rapidly evolving lineages, an unusual phenotype, the nature of which has puzzled microbiologists for decades (176). These organisms have no cell walls; the cell membrane is their outer boundary. They have a number of cytological and biochemical peculiarities, and their genomes are far smaller than normal bacterial genomes (137, 176). Some microbiologists took their unusual phenotypes to mean that mycoplasmas were extremely primitive, not at all related to ordinary bacteria (229); others saw mycoplasmas merely as degener-

ate forms of certain normal bacteria (176). The current consensus among microbiologists, reflected in mycoplasma classification, seems to be that these organisms (with the exception of *Thermoplasma*, an archaeobacterium) constitute a phylogenetically distinct group of highly unusual bacteria that is distantly related to the eubacteria (62, 176).

However, mycoplasmas are not distinctive genealogically. By rRNA measure they are merely gram-positive eubacteria. They and their relatives, as seen above, reside high in the gram-positive tree, as one subline of a particular subgroup in the low-G+C subdivision of that phylum (178, 262, 265). And, as Fig. 15 shows, the mycoplasmas have specific clostridial relatives, for example, *Clostridium innocuum* (262, 265). In other words, mycoplasmas seem unusual not because of a remote phylogenetic position, but because their mode of evolution has for some reason been atypical.

Despite their mundane genealogy, the rRNAs of mycoplasmas are definitely unique (262, 265). As Fig. 15 and Table 18 show, mycoplasma rRNA sequences change more rapidly than do those of normal eubacteria, but more importantly, certain positions in the 16S rRNA sequence that tend to be invariant in composition, what are normally the "phylogenetically uninformative" positions, show significant variation in the mycoplasma sequences (265) (Table 19). Since oligonucleotides representing the regions of highly conserved sequence make a major contribution to  $S_{AB}$  values, the latter tend to drop dramatically in rRNAs that violate these invariances. Table 18 permits a comparison of  $S_{AB}$  values to percent sequence similarities for the mycoplasmas and some of their normal relatives; the  $S_{AB}$  values between *Mycoplasma gallisepticum* and other eubacteria in particular are as low as those between normal eubacteria and archaeobacteria! Varying the normally conserved positions seems characteristic of all rapidly evolving lineages. As mentioned above, a similar disparity between  $S_{AB}$  values and percent sequence similarity holds for the planctomyces (202).

TABLE 18. Percent sequence similarities and  $S_{AB}$  values for 16S rRNAs of mycoplasmas and related eubacteria<sup>a</sup>

Species	<i>M. gallisepticum</i>	<i>M. capricolum</i>	<i>S. citri</i>	<i>Ac. laidlawii</i>	<i>C. innocuum</i>	<i>B. subtilis</i>	<i>H. chlorum</i>	<i>Ar. globiformis</i>	<i>An. nidulans</i>	<i>A. tumefaciens</i>	<i>D. desulfuricans</i>	<i>E. coli</i>
<i>M. gallisepticum</i>		27 <sup>b</sup>	26	19	22	17	14	15	12	09	07	09
<i>M. capricolum</i>	79.5 <sup>c</sup>		47	28	35	30	25	27	20	23	12	17
<i>S. citri</i>	80.5	87.7		25	30	25	23	20	19	22	12	14
<i>Ac. laidlawii</i>	77.4	81.2	79.9		39	28	24	21	17	22	12	15
<i>C. innocuum</i>	76.6	80.8	81.0	82.0		33	27	23	24	22	17	14
<i>B. subtilis</i>	75.0	81.1	79.3	78.6	82.1		41	36	27	26	24	24
<i>H. chlorum</i>	74.5	78.3	78.5	77.9	79.0	83.5		32	24	26	27	28
<i>Ar. globiformis</i>	73.7	78.1	77.2	77.3	79.5	81.0	80.3		24	25	17	24
<i>An. nidulans</i>	73.2	77.0	77.6	76.7	78.4	80.5	81.0	78.6		23	26	24
<i>A. tumefaciens</i>	74.1	77.2	76.5	76.6	77.1	79.1	78.9	77.6	77.7		25	28
<i>D. desulfuricans</i>	72.2	75.0	75.4	75.0	76.8	79.6	80.3	78.7	78.9	80.9		33
<i>E. coli</i>	71.6	75.2	73.9	73.5	75.8	77.5	78.4	77.0	78.0	79.4	79.1	

<sup>a</sup> References 1, 19, 68, 93, 162, 163, 224, 255, 274; sequences for all members of the mycoplasma group except *Mycoplasma capricolum* are unpublished, as is that for *Ar. globiformis*. *M.*, *Mycoplasma*; *S.*, *Spiroplasma*; *Ac.*, *Acholeplasma*; *C.*, *Clostridium*; *B.*, *Bacillus*; *H.*, *Helicobacterium*; *Ar.*, *Arthrobacter*; *An.*, *Anacystis*; *A.*, *Agrobacterium*; *D.*, *Desulfovibrio*; *E.*, *Escherichia*.

<sup>b</sup> Upper right are  $S_{AB}$  values (55).

<sup>c</sup> Lower left are percent similarities; all positions not represented in all sequences have been eliminated from consideration.

### Dynamics of Variation in rRNA

We assume that (naturally occurring) changes in rRNA sequence are "selectively neutral." The assumption is justified on the grounds that translation has to have been one of the earliest functions established in the cell (for the evolution of all of the cell's proteins depends upon its existence) and, once perfected, there should be no reason, no selective constraints, to change it (outside of rare minor adjustments made to accommodate changes in the cell's basic physical parameters, such as optimum growth temperature, intracellular pH, or ionic concentrations). To support this case, one can invoke the facts that rRNA secondary structures show little variation within any of the primary kingdoms (75, 260), implying corresponding functional constancy, and that within a kingdom components can usually be exchanged among different translation systems without destroying function (2, 14, 156). The mycoplasmas are not exceptional in this regard. The physical parameters of their niches are normal; they do not appear to synthesize unusual types of proteins; their rRNAs have normal secondary structures (93); and some of their close relatives, normal eubacteria in all known respects (e.g., *Clostridium innocuum*, *Lactobacillus cateniforme*) share their rRNA sequence idiosyncrasies to some extent (265).

To say that (naturally occurring) changes in rRNA sequence are by and large selectively neutral, does not mean that individual base changes are necessarily so; it is the overall (composite) change that tends to be independent of selection. Changing a single position in a base pair, for example, generally creates a mispair, which would probably be selected against; to change both members of the pair in a way that maintains normal pairing, however, might have negligible selective impact. (More complicated arrangements, tertiary structural and the like, can also be imagined wherein three or more positions would have to be changed to retain proper function.) The individual mutations creating such a composite change, therefore, have to occur simultaneously or at least in fairly rapid succession.

The most important attributes of rRNA sequence variation for our purposes are (i) that recognizable patterns of variation exist, i.e., that the same variations tend to occur in different phylogenetic groups; and (ii) that the frequency with which changes in composition happen can vary widely from position to position within the sequence (262, 265).

Regarding the first point, the pattern of variation at a given position is often the same for most or all eubacterial phyla and sometimes even for archaeobacteria (266). Regarding the second, the rates at which the most and least variable positions change differ by at least two orders of magnitude (265). Both the frequency and pattern of variation seem to correlate strongly with the overall structure of the rRNA molecule and so would seem to be functionally determined. Base pairs at some positions in some helices change frequently; other pairs in these same helices change rarely if at all; some helices are far more variable in composition than

TABLE 19. Number of positions in which a given sequence shows exception to consensus<sup>a</sup>

Species	Consensus sequence			
	1 <sup>b</sup>	2 <sup>c</sup>	3 <sup>d</sup>	4 <sup>e</sup>
Mycoplasmas				
1. <i>M. gallisepticum</i>	66	50	44	99
2. <i>M. capricolum</i>	41	13	24	64
3. <i>S. citri</i>	46	18	23	60
4. <i>Ac. laidlawii</i>	60	20	33	65
5. <i>C. innocuum</i>	23	13	12	39
Normal eubacteria				
6. <i>B. subtilis</i>	6	9	2	22
7. <i>H. chlorum</i>	13	18	3	24
8. <i>Ar. globiformis</i>	27	31	10	41
9. <i>An. nidulans</i>	28			
10. <i>A. tumefaciens</i>	23			
11. <i>D. desulfuricans</i>	17			
12. <i>E. coli</i>	33			

<sup>a</sup> References 19, 68, 93, 163, 224, 255, 274; sequences for all members of the mycoplasma group except *M. capricolum* are unpublished, as is that for *Ar. globiformis*. For genera, see Table 18.

<sup>b</sup> Consensus allowing one exception only at a given position; the values for the species in the mycoplasma group, i.e., no. 1 to 5, are calculated individually from alignments containing that sequence alone (from the mycoplasma group) and no. 6 to 12; the values shown for no. 6 to 12 are therefore averages over the five resulting consensus sequences.

<sup>c</sup> Consensus allowing one exception; alignment contains only gram-positive species, i.e., no. 1 to 8.

<sup>d</sup> Consensus allowing no exceptions; alignment contains species 9 to 12 plus *Leptonema illini* (unpublished data), *Chlorobium vibrioforme* (Weisburg, Ph.D. thesis), *Thermomicrobium roseum* (162) and *Thermotoga maritima* (1); i.e., it contains no gram-positive species.

<sup>e</sup> Same as footnote d except that one exception is allowed in generating the consensus.

others; and the composition of loops tends to be more highly conserved than that of the underlying double-stranded stalks (75, 260, 262, 265).

Why mycoplasma rRNAs are so unusual. Not only do mycoplasmas tend to vary the otherwise conserved positions (i.e., introduce the rare composite changes) in rRNAs more readily than do their normal counterparts, but they even differ significantly from one another in this respect. The tendency to vary conserved positions is far more accentuated in *Mycoplasma gallisepticum*, for example, than in other mycoplasmas (262, 265) (see also Table 20). Moreover, one line of mycoplasmas will make changes in the rRNA sequence that others do not, as though a broad range of possibilities existed from which to choose (262, 265). If their ribosomes are structurally and functionally normal, then functional constraints on the ribosome cannot bring about the evolution of these rRNA idiosyncrasies. One has to consider, therefore, that these idiosyncrasies are not connected to ribosome function/evolution per se; rather, they reflect some general peculiarity of the evolutionary process in mycoplasmas.

Consider the following argument: if changing one of the more conserved positions in an rRNA has to involve (nearly) simultaneous changes elsewhere in the same (or another) molecule, then the occurrence of the overall, composite change is a higher-order function of the organism's (lineage's) mutation rate. For low enough rates such changes occur with a negligible frequency relative to simple (first-order) nucleotide changes, but as the mutation rate increases, the composite changes will increase in frequency relative to the simple ones (265). This means that lineages with low mutation rates have associated with them fields of rRNA variants (from which they derive their evolutionary variability) that are relatively restricted, whereas lines with higher mutation rates draw upon much richer (more varied) fields of variants. Other factors being equal, rRNAs would evolve far more variety in the latter case than in the former.

Although mycoplasma mutation rates have not actually been measured, there are good reasons to suspect them to be abnormally high. In any line of descent mutation rate must be optimized. It cannot be so high that deleterious mutations are created in a significant fraction of an organism's progeny, yet it cannot be so low that the lineage is unable to adapt to fluctuations in its environment or is otherwise unable to compete effectively (265). An upper bound to mutation rate is, therefore, set by an organism's functional genome size. The larger that genome, the harder to replicate it without introducing errors. Consequently, an organism with a small genome could be as stable evolutionarily as organisms with larger genomes even though it had a higher mutation rate (per base pair). Mycoplasma genomes are four to eight times smaller than the eubacterial norm (137). Although mycoplasmas arose from ancestors (clostridia) having normal genomes, and so presumably normal mutation rates, the constraints keeping mutation rates low cease to exist once genome size decreases. One would also expect some mycoplasmas to have elevated mutation rates because they are known to be deficient in certain DNA repair capacities (61) and to lack a DNA polymerase 3'→5' exonuclease activity (150).

**Summary.** In brief, my argument is this: changes in rRNA sequence are for the most part selectively neutral. However, many of these changes are composite and so would appear to involve nearly simultaneous, coordinate alterations in two or more positions in the molecule. Composite change of this type is a higher-order function of a lineage's mutation rate.

Those lines having elevated mutation rates would be expected not only to show the normal types of rRNA variants at higher than normal levels, but also to spawn variants that normally occur at inappreciable levels. Because their genomes are small, mycoplasmas can develop elevated mutation rates, generating rRNA variants not usually associated with normal lineages, which gives a unique richness (variability) to the evolution of their rRNAs.

#### Both Tempo and Mode of Bacterial Evolution Are Reflected in rRNA

A useful, if idealized, model for the rRNA chronometer is a measuring device that comprises a series of counters. The basic, primary counter records the number of (certain) events occurring; it simply measures a distance, a rate  $\times$  time. The others are threshold counters (differing from one another in having progressively higher thresholds). Each one registers nothing until the rate at which the events occur reaches its particular threshold value, after which it too measures rate  $\times$  time. Such a chronometer can measure more than long-term average rates, more than simple distances. It can detect changes in rate and peak rates. rRNA is not the simple uniform-rate chronometer our analyses generally assume. It behaves as a "compound" chronometer in the above sense. As such it should be able to detect whether (but not necessarily when) episodes of rapid evolution have occurred in a lineage's history, which opens the tempo-mode problem to study on the molecular level. Such a rate-sensitive chronometer could also be used to produce phylogenetic trees in which the root is internally delimited (265).

The unusual changes encountered in the rRNAs of mycoplasmas and other rapidly evolving bacterial lineages actually measure the mode of evolution. In a sense they are the mode, for at the molecular level tempo and mode come together; they are different facets of the same process. The mutational changes that are summed to indicate a tempo (evolutionary distance) include the changes that define the quality of the field of variants. The idiosyncratic selectively neutral variants of mycoplasma rRNAs are obviously representative of all variants in the mycoplasma phenotype. Expanding the field of variants, as mycoplasmas appear to do, makes it statistically unavoidable that their evolution be both more rapid than normal and highly atypical.

Macroevolution in mycoplasmas is chronic. As long as mycoplasma genomes remain small, these organisms would seem to be in a chronic state of rapid evolution. Classical macroevolution is episodic, however. Thus, a question remains as to whether bacteria can exhibit episodic rapid evolution, whether macroevolution in its classical form occurs in the microbial world.

The rRNA chronometer in principle should reveal the episodic form of rapid evolution in the same way it does the chronic form, although the former would be harder to detect in that it would leave less of a trace. That rRNA sequence signatures can be constructed for the various bacterial groups implies that unusual sequence changes become fixed during the formation of major taxa. For a group that contains many characterized representatives, the cumulative evolutionary distance within it (i.e., the sum of all branch lengths on the corresponding phylogenetic tree) is very large compared to the evolutionary distance that separates that group's ancestor from the ancestor of some other nearby group. Yet (for these well-characterized groups) the unusual types of change that become fixed during their formation



tend not to occur during the group's evolutionary ramifications. No statistical justification of this point will be attempted at this time, for more rRNA sequence data than now exist are required to make a strong case. However, the indications are there; major bacterial groups come into being through episodic rapid evolution.

**Conditions for true macroevolution among bacteria.** Episodic rapid evolution among bacteria would (according to the above reasoning) require that conditions exist under which bacterial mutation rates can increase but subsequently return to normal. Such conditions may well be those invoked by the classical evolutionist to explain macroevolutionary change (192). (However, the classical conditions now become necessary but insufficient to effect episodic rapid evolution in bacteria.) Environments are of two general types: those to which organisms can become well adapted, and those to which they cannot. Stable or cyclically varying environments can be of the first type. Chaotically varying and "extreme" environments in general are of the second. In the first case an organism's phenotype can (and does) become "fine tuned" to the environment. Organism and niche come into some sort of close and detailed correspondence: nuances of phenotype have selective meaning; additional levels of organization (control) are added; efficiency of function increases; new refinements, details, evolve. Such are the general evolutionary considerations for normal environments. Contrast this to evolution in an unpredictably fluctuating or otherwise extreme environment. Environments of this type stress the organism's physiological responses to the limit. Existence here is a matter of survival under any condition, not of survival of the fittest. In other words, fine-tuning, efficiency, etc., have little to do with evolution in this context. Under these conditions negative selection would, in some senses and some areas, be relaxed. Many genes concerned with fine-tuning would be of little significance. Only the most basic genes, those that make for integrity of the organism and continuity of the lineage, would really count; and the precision or efficiency (i.e., fine-tuning) of their function may not be as selectively significant a concern as it normally is. Under conditions such as these, when many functions become superfluous, the effective genome size is reduced, and some selection is relaxed, a lineage might then be able to sustain an elevated mutation rate. If so, the resultant expansion of the field of variants would mean that unusual phenotypic features necessarily arise in the lineage: some of the rare variations might even be essential to the line's survival in the untoward environment (and so an increased mutation rate would have positive selective value). Were such a line subsequently to adapt to a more stable, compatible environment (or even somehow stabilize in the formerly extreme one), its mutation rate would necessarily return to normal, but the organism would bear the scars of its tumultuous history; its phenotype would be drastically changed and highly unique.

#### Basic Principles of Bacterial Evolution

The above conceptualization of the tempo-mode relationship makes bacterial evolution appear straightforward and understandable. Normal lineages, those having normal mutation rates, do not drastically change their ancestral phenotype. If the environment in which a phenotype arises (first stabilizes) persists, that phenotype will persist, fundamentally unchanged. This is not to say that the original phenotype cannot change to fit a novel environment (without increase in mutation rate and so on), but it is to say that the

kinds of new environments to which it is capable of adapting in this way do not radically alter the ancestral phenotype. On the other hand, all drastic (broad-ranging) changes in ancestral phenotype necessarily result from increased mutation rates, which tend to occur under unusual, drastic environmental conditions, when selection is relaxed in ways that allow the mutation rate to rise.

In a general sense, then, the course of bacterial evolution is relatively simple to chart by means of a macromolecular chronometer, a purely genotypic measure. Lineages represented by the shorter branches on a phylogenetic tree (i.e., the ones least distant from the tree's root) are slowly evolving and retain proportionately more of the ancestral phenotype. Lineages represented by the longer branches, on the other hand, are rapidly evolving and necessarily retain far fewer ancestral characteristics. Moreover, a molecular chronometer will show structural idiosyncrasies in the latter case, the "scars" of the lineage's bout of rapid evolution.

#### Implications for Bacterial Taxonomy

If future findings support the above conclusions and speculations, then it should be possible to construct a bacterial taxonomy based upon naturally defined categories. Groups that arise through macroevolution are self-defining; they are recognizably unique, distinct both phenotypically and genotypically (in terms of molecular chronometers). Many higher bacterial taxa, i.e., phyla and their major subdivisions, show this characteristic.

In addition to naturally demarcated categories, two other requirements for a natural taxonomy are: (i) a means of determining relationships among the categories, which, of course, is given by the topology of a phylogenetic tree; and (ii) a means of naturally defining taxonomic rank, one that is not completely dependent upon tree topology. There is no reason a priori that a taxon of higher rank, a class, for example, cannot be included in one of lower rank, e.g., a family. Indeed, the mycoplasmas may be a case in point. The question is whether the microbiologist can accept such natural "inversions" of rank as taxonomically valid or whether he will insist upon an arbitrarily defined taxonomy that does not contain them and so appears orderly to him.

Macroevolutionary episodes can probably be classified by the degree of their severity. If the rRNA chronometer in essence consists of a series of counters having progressively higher thresholds (i.e., the model used above), then the highest-threshold counter activated by a macroevolutionary episode defines the severity of the episode, which is a de facto definition of taxonomic rank. Without further constraints, however, we are again left with the possibility of a jumbled taxonomic hierarchy, for the timing and severity of macroevolutionary episodes would seem to be unrelated in unrelated lineages.

It is obvious that metazoan taxonomy exhibits more order than is inherent in the natural system just described. The animal phyla seem all to have arisen at about the same time, somewhat less than 1 billion years ago; the eucaryotic kingdoms arose similarly at an earlier stage (196). Evolution tends to follow two rules: (i) the higher its taxonomic rank, the further back in time the taxon arises; and (ii) major taxa of the same rank tend to arise in the same era, giving rise to evolutionary radiations. All this may seem self-evident to some, but it is not, particularly in the bacterial world where potential ancestral phenotypes seem to persist for all taxa up to the kingdom level. The explanation for this unaccountable time ordering of the actual taxonomic hierarchy may lie

outside of biology per se, in the evolutionary history of the planet. Geologists are coming to the conclusion that the history of earth and the other planets, moons, etc., is written in terms of catastrophes, relatively sudden, chaotic global changes (3). The frequency and severity of such changes increase as one looks backward in time. (Unfortunately, the record of their happening erodes with time, particularly on earth.) The intense bombardment by meteors and similar objects, which changed the face of the moon (and presumably affected earth as well) prior to 3.9 billion years ago, is one example (85). Another example, more germane to biology, might be the relatively sudden (on an evolutionary time scale, that is) rise in atmospheric oxygen concentration. Such precipitate shifts in global parameters would be capable of triggering synchronous macroevolutionary episodes in various lineages (3). Because the more severe "catastrophes" tend to be the earlier ones, the taxa of higher rank would tend to form earlier than those of lower rank.

The genealogical history of bacteria beginning to be revealed by molecular chronometers will provide ample data to develop or discard such a view of evolutionary relationships and their taxonomic implications.

### EVOLUTION OF THE TWO PROCARYOTIC PHENOTYPES

The enormous value of bacterial phylogeny as a classification system, a predictive and organizational framework, is easier to appreciate than its value as an historical account and source of evolutionary insights. This is because our evolutionary perspective has to this point in time been focused narrowly on metazoa and their fossils. This point of view has inevitably overemphasized morphology, concentrating on subtle but superficial differences among complex forms. At the same time it has underemphasized the "metabolic" aspect of evolution; it embodies little feeling for biochemistry and energy flow. Since it has been confined to a relatively recent (and so, in the grand scope, uninteresting) period in earth history, well after formation of the oxygen atmosphere (180, 228), our present view incorporates in a minor way only the dynamism of the evolving earth and the close relationship between its physical and biological evolution. Overall, this view is a static one, in which evolution is not a process but rather "a procession of forms," as Whitehead put it (237).

In contrast, microbial evolution is essentially metabolic, fundamentally biochemical. It spans the bulk of our planet's history and is intimately tied thereto. As the base of the food chains, microbial metabolic patterns bear a straightforward and (ultimately) understandable relationship to the planet's geochemistry. Thus, bacterial evolution is no simple extrapolation of metazoan evolution—more of the same and, lacking fossils, harder to study. Microbial evolution is a different story, told in a different way, covering a different (more extensive) period of earth history, a story that is simpler, more readily interpretable, and more informative of the planet's physical course.

#### Archaeobacterial Evolution

**Ancestral phenotype.** It would appear that the ancestral archaeobacterium was an extremely thermophilic anaerobe that probably derived its energy from the reduction of sulfur. Two lines of evidence support such a conclusion. The first is the widespread distribution of the extreme thermophilic phenotype among the archaeobacteria. Of the three basic

archaeobacterial phenotypes, it is, as we have seen, the only one that occurs on both major branches of the archaeobacterial tree. *Thermococcus celer*, on the methanogen branch, and *Pyrodictium occultum*, on the extreme thermophile branch, are both typical sulfur-metabolizing thermophiles, thriving anaerobically in hot spring environments. (*Pyrodictium* holds the current record for highest optimum growth temperature of any organism, 105°C [212], while *Thermococcus* is a common inhabitant of marine hot springs [280].) The second is that the extreme thermophile phenotype is the only one to meet the tempo-mode criterion for being ancestral; the evolutionary distance between *Thermococcus* and *Pyrodictium* is remarkably short, about 18%, appreciably shorter than the shortest distances, 24 to 25%, that separate any methanogen (relatives of *T. celer*) from *P. occultum* or its relatives, as can be seen in Table 17. From these facts, and the fact that the two extreme thermophiles are about equally distant from various eubacterial or eucaryotic outgroup sequences (263), it follows that both lineages are slowly evolving and, therefore, have retained more common ancestral characteristics than have the other archaeobacterial phenotypes.

Two genera, *Sulfolobus* and *Thermoplasma*, are atypical sulfur-dependent thermophiles in having evolved the capacity to utilize oxygen (215), almost certainly a derived characteristic. It is interesting, therefore, that both represent relatively rapidly evolving lineages. The lineage of *Sulfolobus* is the most rapidly evolving on its branch of the archaeobacterial tree, while that of *Thermoplasma* is perhaps the most rapidly evolving of all archaeobacterial lineages (263).

**Evolution of methanogenic and halophilic phenotypes.** Since these two phenotypes are significantly further from the extreme thermophile cluster (by rRNA measure) than is their relative *Thermococcus*, the methanogenic and halophilic lineages would seem to have undergone macroevolution. Presumably such an episode was associated with the transition from an ancestral thermophilic sulfur-metabolizing phenotype to a methanogenic one. Given such an evolutionary progression, one has the interesting question of how thermophilic sulfur metabolism can change into methanogenesis and what global conditions might have favored, i.e., brought about such a transition. (Very recently K. O. Stetter [personal communication] has isolated what may be a missing link in such a transition. The organism is a novel archaeobacterial phenotype; it grows anaerobically and reduces sulfate. It contains several of the cofactors characteristic of the methanogens, but lacks the all-important factor coenzyme M, which is involved in the terminal step of methane production [238]. Nevertheless, the organism does produce methane in minute amounts, as does the eubacterial sulfate reducer *Desulfovibrio desulfuricans* [173]. A preliminary and unpublished partial 16S rRNA sequence shows its lineage to arise from the methanogen branch of the archaeobacteria between the *Thermococcus* and *Methanococcus* lineages in Fig. 13.)

Within the methanogens per se a second round of rapid evolution seems to have occurred, involving the branch that leads to the *Methanomicrobiales*. Note in Table 17 that sequence distances for members of the *Methanomicrobiales* are greater than corresponding distances for both of the other methanogen phyla. This can be seen in convincing detail in the analysis of Table 20, which gives sequence distances of various archaeobacterial species from various consensus sequences.

Although the most spectacular change during this second

TABLE 20. Sequence distances (in percentages) of archaeobacterial rRNAs from various consensus sequences<sup>a</sup>

Sequence from:	Consensus sequence			
	1 <sup>b</sup>	2 <sup>c</sup>	3 <sup>d</sup>	4 <sup>e</sup>
<i>Mc. vannielii</i>	— <sup>f</sup>	—	10.0	16.0
<i>M. formicicum</i>	—	—	9.6	15.4
<i>Ms. hungatei</i>	—	10.0	13.9	19.3
<i>Mr. sp. strain WH-1</i>	6.5	10.5	13.0	18.6
Extreme halophiles <sup>g</sup>	8.9	11.6	15.6	20.7
<i>T. acidophilum</i>	13.1	13.0	15.7	21.6
<i>Tc. celer</i>	5.1	—	—	10.4
<i>S. solfataricus</i>	12.3	12.1	5.5	6.7
<i>Tp. tenax</i>	11.5	10.4	—	—
<i>D. mobilis</i>	8.9	8.3	—	—
<i>P. occultum</i>	9.3	8.3	1.1	—

<sup>a</sup> Only those positions represented in all positions in 16S rRNA alignment are used in calculation. References 72, 91, 96, 124–126, 158, and 275. Sequences for *Methanosarcina* sp. strain WH-1, *T. acidophilum*, *Tc. celer*, *D. mobilis*, and *P. occultum* are unpublished. *Mc.*, *Methanococcus*; *M.*, *Methanobacterium*; *Ms.*, *Methanospirillum*; *Mr.*, *Methanosarcina*; *T.*, *Thermoplasma*; *Tc.*, *Thermococcus*; *S.*, *Sulfolobus*; *Tp.*, *Thermoproteus*; *D.*, *Desulfurococcus*; *P.*, *Pyrodicticum*.

<sup>b</sup> Consensus based upon *Mc. vannielii*, *M. formicicum*, and *Ms. hungatei*.

<sup>c</sup> Consensus based upon *Mc. vannielii*, *M. formicicum*, and *Tc. celer*.

<sup>d</sup> Consensus based upon *Tc. celer*, *Tp. tenax*, and *D. mobilis*.

<sup>e</sup> Consensus based upon *Tp. tenax*, *D. mobilis*, and *P. occultum*.

<sup>f</sup> —, Sequence used to generate consensus.

<sup>g</sup> Values are averages of the three published halophile sequences and are within 7% of one another.

saltation was the conversion of an anaerobic methanogen into an aerobic extreme halophile, profound changes also affected the methanogens themselves in this particular lineage. Some of the *Methanomicrobiales* became halophilic, some even alkaliphilic (13; Mathrani et al., Abstr. Annu. Meet. Am. Soc. Microbiol. 1985). *Methanosarcina* and its relatives learned to produce methane from acetate or methyl amines; they are, as noted, the only methanogens possessing cytochrome *b* or *c* or both (112). The extreme halophiles, which require small amounts of oxygen to synthesize their carotenoids (79), may represent an evolutionary response to the onset of aerobic conditions in the hydrosphere, about 1.5 billion years ago (228).

**Archaeobacterial ribosomes and their evolutionary implications.** The archaeobacteria are unique among the three urkingdoms in that variation in their ribosome type occurs. Cammarano and co-workers (21) have measured the protein content of ribosomal subunits by buoyant density centrifugation and find that the molecular weight of protein associated with the small subunit in the extreme thermophiles is  $0.64 \times 10^6$  to  $0.66 \times 10^6$  but in the extreme halophiles and two of the methanogen phyla it is only half this, i.e.,  $0.31 \times 10^6$  to  $0.32 \times 10^6$ . In the remaining methanogen group, the *Methanococcales*, protein contents of the small subunit have an intermediate value,  $0.52 \times 10^6$  daltons. A similar situation obtains for the large subunit; there the protein contents in the extreme thermophiles and *Methanococcus* are in the range of  $0.77 \times 10^6$  to  $0.97 \times 10^6$  daltons, but in the remaining methanogens and the extreme halophiles they are much less, i.e.,  $0.51 \times 10^6$  to  $0.57 \times 10^6$  daltons (21). (The ribosomes of *Thermoplasma* are also high in protein; the small and large subunits contain  $0.61 \times 10^6$  and  $0.78 \times 10^6$  daltons of protein, respectively [21].) Since rRNA sizes are very nearly the same for all archaeobacteria, this approximate twofold variation in protein content must make some change in the size of the ribosomal subunits, and electron micrographs bear this out. The large ribosomal subunit from the

extreme thermophiles is larger than those seen in methanogens; it has several protrusions that its methanogenic (and halophilic) counterparts do not (81). However, some of the large ribosomal subunits from *Methanococcus vannielii* have these protrusions as well, while the remainder are typical of the other methanogens (and extreme halophiles) (218, 219).

The molecular basis for these large differences in ribosomal protein content is not understood. There is no indication that they are connected to structural differences in the rRNAs (Woese, unpublished analysis). The excess proteins are probably not, therefore, attached directly to rRNA. It also seems unlikely that the drastic disparities in protein content reflect any significant difference in ribosome function in the two classes.

When these differences in ribosome type were first discovered (as shape differences in electron micrographs), they were interpreted to mean that the archaeobacteria were not a valid taxon; the extreme thermophiles and the methanogens (and their relatives) constituted separate urkingdoms, the former having "a close relationship to eukaryotes" (115). (For similar reasons, the extreme halophiles were subsequently extracted from the methanogen branch to constitute another new kingdom, the "photocytes," which ostensibly was specifically related to eubacteria [114].) Now that ribosome morphologies have been more thoroughly investigated, it is apparent that ribosome shapes (protein contents) constitute more a spectrum of types than two clear-cut classes (114, 115, 218, 219). In any case, were separate kingdoms to be defined along these lines, one group of methanogens, the *Methanococcales*, would end up in a different kingdom (218, 219) than the others—which is absurd! As taxonomists well know and have repeatedly stated, a small number of (ill-defined) characters is an unreliable basis upon which to define taxa.

**Summary.** Archaeobacterial evolution can be simply understood in terms of an aboriginal anaerobic thermophilic sulfur-metabolizing phenotype that remained pure in one of the urkingdom's two main lineages, but gave rise in the other, through several macroevolution episodes, initially to methanogenic metabolism and then to an altered, more versatile (acetate- or methylamine-utilizing) form of methanogenesis, to halophilic methanogens, and ultimately to the aerobic (nonmethanogenic) extreme halophiles.

### Eubacterial Evolution

Eubacterial history seems less straight-forward than archaeobacterial history. The most prominent phenotypic characteristics of the eubacterial tree are its metabolic diversity and the widespread distribution of anaerobic, photosynthetic, and thermophilic phenotypes. Aerobic and anaerobic groupings stand in sharp contrast to one another. Whereas the various anaerobic phenotypes more often than not form phylogenetically deep, extensive groupings, such is not the case for their aerobic counterparts (56, 206). A good example is the comparison between the clostridia (a remarkably deep anaerobic phylogenetic unit) and *Bacillus* (a phylogenetically much shallower collection of aerobes) (56). Thus, although aerobic phenotypes evolved a number of times, their lack of phylogenetic depth suggests that they are all relatively recent in origin.

**Photosynthesis.** In slightly different versions, photosynthesis appears in at least half of the eubacterial phyla. The purple, blue-green, gram-positive (*Heliobacterium*), and green sulfur lineages stem from the same general area on one of the two major branches of the eubacterial tree (Fig. 11).



FIG. 16. Electron micrograph of a square extreme halophile (216); figure courtesy of W. Stoeckenius.

The remaining photosynthetic type, the green non-sulfur bacteria, represents a significantly deeper branching in the tree (162). Within some phyla, e.g., the purple bacteria, the photosynthetic phenotype is intimately intermixed with nonphotosynthetic phenotypes (162, 267–269). Given the complexity of the photosynthetic apparatus, it seems unlikely that the process has evolved more than once within the eubacteria, and so its origin is deep in the eubacterial tree, possibly at the stage of the common eubacterial ancestor.

**Thermophilia.** It can be even more compellingly argued that thermophilia is an ancestral eubacterial characteristic. All deeper branchings in the eubacterial tree involve either predominantly or exclusively thermophilic groups (1): the *Thermotoga* lineage forms one side of the deepest known branching in the eubacterial line. (The organism so far has only one known [unnamed] relative, strain H21 of K. O. Stetter, again a thermophile [Stetter and Woese, unpublished data].) As mentioned above, the unusual thermophile *Thermodesulfotobacterium commune* (119) represents another very deep branching. A third is formed by *Thermomicrobium* and its relatives (the green non-sulfur bacteria), again a predominantly thermophilic group (94, 162, 170). All of these thermophilic lineages are also relatively slowly evolving, especially *Thermotoga*, and should, by the tempo-mode criterion, be primitive in type. The only known mesophile to branch deeply in the eubacterial tree is *Herpetosiphon* (a relative of *Chloroflexus* and *Thermomicrobium*). However, this organism is the product of a rapid evolution (162) and so should not resemble the ancestral phenotype to the extent that its thermophilic relatives do. Thermophiles are also found in the "mesophilic section" of the eubacterial tree, e.g., *Bacillus stearothermophilus* and *Thermus aquaticus*. In almost all of these cases, too, the thermophilic lineages are at least as slowly evolving as their specific mesophilic relatives. The evidence clearly points to a thermophilic origin of the eubacteria.

One could similarly, but less convincingly, argue an autotrophic ancestry for the eubacteria, for autotrophy is also widely distributed among eubacteria.

**Conclusion.** The case for a photosynthetic, thermophilic, or autotrophic eubacterial ancestor can never be proven in

the strictest sense. But that is not what is important. Because we can now approach microbial phylogeny experimentally and can construct a microbial tree, we are in a position to weed out former incorrect notions and develop new and more detailed ones more soundly based upon experimental evidence. We are on the threshold of greatly expanding our understanding of bacteria and their relationships to one another—and proceeding from there to a reconstruction of the history of life on this planet.

#### Differences between Archaeobacteria and Eubacteria

The persistent influence of the procaryote-eucaryote dogma prevents many biologists from appreciating the importance of treating archaeobacteria as distinct from eubacteria: "They may be very different, but they are still procaryotes" is an attitude one still hears repeatedly. Any such feeling inhibits understanding of the eubacterial-archaeobacterial relationship. Even the fact that the names of the two groups bear the common suffix "-bacteria" acts to do so. Therefore, it is important at this juncture to stress the differences between the two procaryotic types.

**Cytological and physiological differences.** Eubacteria and archaeobacteria are perceived as cytologically similar mainly because both are simpler than and do not resemble the more complexly structured eucaryotic cell. Within their "similarity," however, lie significant differences in cell architecture and metabolic patterns and remarkable differences in evolutionary behavior. Cells of many extreme halophiles are very thin, flat, and straight sided, with precisely square corners (216, 230); Fig. 16 shows a remarkable example. Flat pseudo-geometric shapes have also been noted occasionally among the methanogens (242). Since flat angular geometries are foreign to eubacteria, basic differences between archaeobacteria and eubacteria in cell architecture are implied.

Methanogenesis involves a variety of coenzymes unique to that process, and at least one unusual coenzyme is somehow associated with sulfur metabolism in the extreme thermophiles (Fig. 12). It is hard to believe that such uniqueness is entirely confined to specialized biochemistries

such as methanogenesis; some of it must carry over into the cell's general metabolism. In other words, the unique coenzymes imply significant differences between archaeobacterial and eubacterial intermediary metabolisms. The area of archaeobacterial intermediary metabolism is ripe for further study (57).

As mentioned above, archaeobacterial membranes may be somewhat different structurally from their eubacterial counterparts, for most of them contain a significant level of diglycerol tetraethers, something unprecedented among eubacteria and eucaryotes (117, 118).

Genome organization and control are significantly different in the two classes of procaryotes; one has the feeling that, if more were known, their difference would seem even more profound. Some archaeobacterial genomes contain families of repeat sequence DNA, not a eubacterial characteristic (39, 182). Although the sequences of many eubacterial genes are known, introns have not been reported. Yet among the few sequenced archaeobacterial genes, a number contain introns: the majority of sequenced tRNA genes from *Sulfolobus solfataricus* contain introns (100; B. P. Kaine, manuscript in preparation), as does one tRNA gene from an extreme halophile (30) and an archaeobacterial 23S rRNA gene (107). Nothing so far indicates that archaeobacteria use the well-known eubacterial types of gene regulation mechanisms. One would think that, if they did, some evidence for their doing so would by now exist.

**Niches.** The two types of procaryotes tend to inhabit different types of environments. Archaeobacteria prefer high-temperature niches. One side of the archaeobacterial tree appears to consist only of thermophilic species. Methanogens that grow at high temperatures are also fairly common (214, 278). No major well-characterized eubacterial group is known to be exclusively thermophilic, although the *Thermotoga* side of the eubacterial tree, represented so far by only two species, remains potentially so (1, 89). Most eubacterial groups are not even predominantly thermophilic. Eubacteria, on the other hand, adapt more readily than archaeobacteria to the myriad low-temperature niches, wherein they predominate.

While eubacteria readily adapt to aerobic conditions, archaeobacteria seem to have difficulty in doing so. Among archaeobacteria, even facultative aerobes are relatively uncommon; there are no obligate aerobes (79). The evolution of oxygen utilization in archaeobacteria appears to be associated with episodes of rapid evolution (see above), which is not the case with (at least some) eubacteria; *Bacillus*, for example, although aerobic, represents one of the most slowly evolving of eubacterial lines by sequence distance measures (56, 163, 265).

**Metabolic versatility.** The evolutionary differences between the two classes of procaryotes manifest themselves in several ways. One is an ill-defined evolutionary quality that, for want of a better term, will be called "metabolic versatility." Eubacteria are remarkable for the variety of their metabolisms, both the overall variety and the extent to which variation occurs within some of the phyla and subdivisions. By comparison, archaeobacteria are metabolically monotonous. Little variation exists within the three basic archaeobacterial phenotypes; the utilization of acetate and methyl amines by one group of methanogens (discussed above) is an example of the sort that does occur. While photosynthesis is a prevalent, perhaps ancestral, theme among eubacteria, no archaeobacteria are completely photosynthetic; the extreme halophiles, however, have acquired the capacity to use light to run various molecular pumps

(122, 217). It is remarkable how little metabolic convergence the two procaryotic groups have shown over their several billion years of coexistence.

**Molecular plasticity.** Paradoxically, the archaeobacteria show more variety than eubacteria do in another evolutionary parameter, a quality we will call "molecular plasticity," i.e., the variations in molecular design of a given function within a group. That the archaeobacteria are exceptional in this respect has been apparent from the outset (250). Examples of the group's molecular plasticity (mentioned above) are seen in 5S rRNA secondary structure (53, 133, 208), DNA-dependent RNA polymerase subunit patterns (187, 282), ribosome protein contents (21), the extent of post-transcription modification of bases in rRNAs and tRNAs (70, 71, 73, 259), cell wall structure (101, 109, 111), types of coenzymes (238), and antibiotic sensitivity (22). In all of these cases archaeobacteria present at least two distinctly different types, more often a spectrum of types, while eubacteria present a relatively monotonous picture of uniformity (250).

**Rates of evolution.** The two classes of procaryotes seem to be evolving at generally different rates. Evidence suggesting this is seen in the phylogenetic breadth of the two groups, i.e., the sequence distance separating the most slowly evolving representatives of the deepest branches in each tree. The phylogenetic breadth of the archaeobacteria (i.e., the distance between the extremely thermophilic representatives on its two main branches) is under 20%; see Table 17. For eubacteria the comparable distance (between *Thermotoga* and *Bacillus* [1]) is slightly under 30%. Since the root of the universal tree is not yet known, it is possible that all this means is that the archaeobacteria are not as old a group as are the eubacteria (attributing the smaller sequence distances in this case to shorter time of evolution, rather than a slower rate). Such an assumption would require the root of the universal tree to be placed relatively high on the eubacterial branch, which then makes the archaeobacteria and eucaryotes quite specific relatives of one another and the eucaryotes a very rapidly evolving line of descent, assumptions I find intuitively unappealing.

**Summary.** This discussion of eubacterial-archaeobacterial differences amounts briefly to this: although ostensibly similar cytologically, the two groups of procaryotes are significantly different in cellular makeup and in their modes of evolution. They tend to inhabit basically different niches. They appear to evolve at different rates. They show differences in two evolutionary parameters, metabolic versatility and molecular plasticity. And, over several billion years of coexistence they have shown little or no tendency to converge in phenotype. It is as though the two have evolved in different worlds.

#### Why Are Archaeobacteria and Eubacteria So Different?

How are we to understand these substantial differences between the two types of procaryotes? Are they a matter of different environments during the early stages of evolution in the two lines? Are basic organizational differences in the two types of cells being reflected in different evolutionary proclivities? We are in no position to answer such questions now. It is not simply that facts are lacking; key concepts seem to be missing too. The best we can do now is to ask questions that we hope will lead us to experiments that provide the required insights.

One key question seems to be whether archaeobacteria are significantly older than eubacteria (or vice versa). If

archaeobacteria already existed before the (most recent) common ancestor of all eubacteria arose, then the basic archaeobacterial phenotype would probably have evolved to suit a global environment very different from that in which eubacteria later arose, making the two cell types basically and perhaps irrevocably dissimilar.

A second key question is whether the archaeobacterial ancestor was more primitive than its eubacterial counterpart, and if so, in what ways. A more rudimentary, less highly integrated ancestor would conceivably have a broader spectrum of potential phenotypes into which it might evolve than would a more highly integrated, more constrained and "advanced" ancestor. An archaeobacterial ancestor of this type would explain the greater molecular plasticity of the group.

Finally, one wants to know the relationship between the common ancestors of both groups (and their relationship to the eucaryotic ancestor). What sort of cell, entity, or system gave rise to the three urkingdoms? It is unlikely that we will ever know the full answer here; however, it is something about which we will soon be able to infer a great deal more than we can now.

#### Primitiveness of Archaeobacteria

In the past, discussion of what was and was not a primitive characteristic was more or less fruitless. (What appeared primitive to some often turned out merely to be degenerate, e.g., the mycoplasmas.) This is no longer the case, for molecular chronometers provide a nearly certain definition of "primitive." In a group of homologous molecular functions the one whose sequence is closest to that of the common ancestral version is necessarily the most primitive. Provided the position of its root can be fixed, a sequence tree then decides the issue.

Current evidence tentatively suggests that archaeobacteria are probably the more primitive of the two procaryotes, in three senses of the word: (i) the group as a whole is older than the eubacteria; (ii) their common ancestor was a more primitive type of entity than the eubacterial common ancestor; and (iii) since they evolve at a slower rate than do the eubacteria, archaeobacteria today remain more primitive (more ancestral in type) than eubacteria.

The most convincing evidence comes, of course, from the rRNA chronometer. Since the archaeobacterial 16S rRNA is closer in sequence to both its eubacterial and eucaryotic counterparts than these two are to one another, the archaeobacterial version of the molecule must be closer to the common ancestral version than is one or both of the other versions (72). Placing a root on the (unrooted) universal tree anywhere within a zone that includes the archaeobacterial branch and a fair segment on both of the other main branches (which can be visualized on Fig. 4) would make the archaeobacteria more primitive than both of the others.

A similar, but weaker, conclusion can be drawn from the relationship among the DNA-dependent RNA polymerases in the three urkingdoms. By serological cross-reactivity, the archaeobacterial polymerases appear closer to their eubacterial and eucaryotic counterparts than these two types are to one another (90, 187). (Sequencing studies now in progress should soon permit more definite conclusions.) This tendency for archaeobacteria to be closer to both of the other groups than they are to one another also carries over (in a qualitative sense) to the general phenotype (250). In a few years, when sequence evidence is available for a significant number of different molecular functions, it may be possible

to say in a quantitative way that the general archaeobacterial phenotype is more primitive than at least one of the others. If so, the question of archaeobacterial primitiveness is half answered.

Knowing the root of the universal tree (the ancestral point) would automatically determine which of the three phenotypes is the most primitive. Conventional wisdom holds that the root of the universal tree cannot be determined, because no outgroup exists by which to position it. However, the root of the universal tree can be determined, in principle if not in practice. What is required is a gene that has duplicated in the common ancestor state (as pointed out by M. Dayhoff long ago). If two (functionally distinct) versions of such a gene fulfilled certain technical requirements, they could then be used in effect to determine relative rates of evolution within each lineage, thereby fixing the tree's root. A practical system for doing this does not yet exist.

#### When Did Procaryotes Evolve?

The general nature of the archaeobacterial phenotype strongly implies the nature of the environment in which it arose. Given the widespread distribution of thermophilic species (and their universal occurrence on the extreme thermophile branch), it seems impossible that these organisms arose from a mesophilic ancestor. The ancestral archaeobacterium was a thermophile, probably growing at temperatures near the present boiling point of water (see above discussion). This makes it likely that the archaeobacteria arose when the ambient temperature of the planet was high, i.e., within the first billion years or so of earth history (43, 271). The ancestral archaeobacterial environment seems also to have been highly reducing, for most archaeobacteria today are fastidious anaerobes, again implicating rather early stages in earth history, when both hydrosphere and atmosphere would have been reducing (228).

Since stromatolites existed at least 3.5 billion years ago (231), eubacteria such as *Chloroflexus* almost certainly existed at that time (162). Given the distribution of thermophilia among eubacteria (see above), these organisms too appear to be of ancient and thermophilic origin. However, in eubacteria thermophilia is not as extreme as in archaeobacteria; finding archaeobacteria that grow at the boiling point of water is now common, but thermophilic eubacteria growing above 90°C are rare. The ancestral eubacterium might then have arisen later than the archaeobacteria, when the planet was somewhat cooler (43, 271).

The indications are there: archaeobacteria seem an ancient and primitive phenotype, more so than the eubacteria. The experiments that would convincingly establish the point are apparent and some are now being done. Eucaryotes remain the puzzle. Were they too of thermophilic origin? How do they fit into this scenario developed for the bacteria?

#### THE UNIVERSAL ANCESTOR

All questions concerning relationships among the three urkingdoms and the general course of evolution in each ultimately turn upon the nature of their common ancestor. The nature of this universal ancestor is, in my opinion, probably the most important, and definitely the least recognized, major question in biology today. As we shall see, the universal ancestor may have been a kind of entity outside of our direct experience. Even to begin considering it, we have to question concepts generally taken for granted.

Consider how information is organized in various living systems. In procaryotes the bulk of the information in the cell occurs in very long, contiguous strings of genes; procaryotes could then be called "genomic" organisms. Lower eucaryotes are also genomic, but the higher ones, metazoa with complex development systems, higher nervous function, and social structures, should probably be designated "supragenomic" entities, for much of the information they contain lies in structures above the level of the genome. Entities simpler and more primitive than genomic ones must also have existed. An organism of this type could have had a genotype and phenotype (i.e., information stored in a quiescent [replicative] form in one class of molecules that was also manifested in an active [functional] form in another), but its genes would for the most part have been physically separate units; they would not be organized into large contiguous linear arrays. These less organized systems would be called "genetic" but not genomic entities. (The reason for distinguishing genetic from genomic entities will become apparent as we proceed.) At a still more primitive level, entities can be imagined in which the genotype and phenotype do not exist in the sense we know them. Instead the storage/replicative and the active/functional forms of information both reside in the same class of molecule, possibly in different configurations of a given molecule. This would be the stage of "nucleic acid life" (245-247), where translation as we know it has not yet evolved and nucleic acids have both genetic and enzymatic functions. The idea that there could have been such an early stage has recently become quite popular with the experimental demonstration that RNAs can have catalytic properties (26, 232). Still more primitive stages can be imagined in which the bulk of the information in the system is not in macromolecular primary structures, but is contained in autocatalytic/metabolic networks.

### Progenotes

The progenote is a theoretical construct, an entity that, by definition, has a rudimentary, imprecise linkage between its genotype and phenotype (251, 256). (Extant organisms, which have precise, accurate links between genotype and phenotype, are then genotes.) The certainty that progenotes existed at some early stage in evolution follows from the nature of the translation apparatus. Translation is accomplished by a multicomponent (multigene) mechanism that includes on the order of 100 different macromolecular species, far too complex a system to have arisen initially fully formed. Like the radio, the automobile, and similar devices, translation had to evolve through stages, from a much more rudimentary mechanism to the present precisely functioning one (244-246, 248, 256). Its aboriginal forms had fewer components and, consequently, must have functioned less accurately than their modern counterparts. There seems no alternative to the conclusion that the progenote existed at some stage early in evolution.

**Characteristics of the progenote.** The limitations of its rudimentary translation mechanism ensure that the progenote was a highly unique entity, unlike any life found today. Without today's level of accuracy in translation, proteins of normal size could not have been synthesized without introducing (many) errors. This means the progenote could neither have had nor have evolved "modern" proteins (244). Its proteins would have been small or of nonunique sequence or both. (A collection of polypeptides all different from one another, but each an approximate translation of the same

genetic sequence, is known as a "statistical protein" [244].) As a consequence the progenote's enzymes would not be as accurate and specific as their modern counterparts. This in turn would delimit the kinds of control mechanisms, the definition and number of states the system possessed, and so on. Biological specificity at the progenote stage had to have been generally lower than now exists (251).

Replicating a genome places a tremendous burden of accuracy on a cell. To reproduce a string of nucleic acid thousands of genes long without introducing significant error requires an extremely precise mechanism, which today involves a number of separate activities (proof-reading functions, error correcting systems, and so on), most of which utilize large proteins. Such an extensive, complex, and precise system would not be found in a progenote, which means that the progenote could not have carried (could not have accurately replicated) the number of genes found in modern cells (251). A factor of 10 drop in the accuracy means a proportionate reduction in the length of the genome. The progenote reasonably had error rates two, or even three, orders of magnitude greater than found in cells today. (A factor of 100 drop in accuracy would leave the mistake rate in the range of 1 part per million [monomer units introduced], which is still impressively accurate.) One therefore wonders how progenotes could have carried a sufficient number of (different) genes to make them even minimally functional cells.

This apparent paradox can be resolved by making the progenote a genetic, not a genomic, entity. Genes would then be disjoint, and they could have existed in high copy numbers, in which case an appropriately simple mechanism can be imagined that would detect errors in individual genes and selectively eliminate (not correct) the flawed ones (251). As a genetic entity, the progenote could, in spite of a relatively very error-prone gene replication process, carry a reasonable number of genes. Given disjoint genes that might assume functional configurations (251), it is likely that the informational macromolecule at this stage was RNA, the functional form of nucleic acid today, not DNA (251).

### Was the Universal Ancestor a Progenote?

In principle the universal ancestor could have resembled any one of the three major types of extant organisms. It also could have in essence been a collage of all three, or have been very unlike any of them. I will argue that the last alternative is the correct one and that the universal ancestor was a progenote.

The evolution that transformed the universal ancestor into the individual ancestors of each of the three primary kingdoms was of a unique quality. Sequence distances between kingdoms (Fig. 4) seem large compared to the distances within kingdoms, this despite the fact that the bulk of evolutionary time has involved evolution within the kingdoms. (The existence of 3.5 billion-year-old stromatolites [231] implies the existence of photosynthetic bacteria at that time, and so the existence of the common eubacterial ancestor even earlier.) Therefore, the long sequence distances do not correspond to long times. The transition from the universal ancestor to the ancestors of each of the primary kingdoms had to have taken less than 1 billion years, perhaps far less if an appreciable fraction of earth's first billion years involved evolutionary stages that preceded that of the universal ancestor. It would seem that the tempo of evolution at the time of the universal ancestor was very high.



The types of phenotypic changes that accompanied the formation of the three primary kingdoms are of a special nature. General differences in cell architecture among the three groups are remarkable, as are their differences in intermediary metabolism, and each kingdom seems to have its own unique version of every fundamental cellular function: translation, transcription, genome replication and control, and so on. The kind of variation that subsequently occurred within each of the kingdoms is minor by comparison. Thus the mode of evolution accompanying the transition from the universal ancestor is unusual; far more novelty arose during formation of the primary kingdoms than during the subsequent evolutionary course in any one of them.

It is hard to avoid concluding that the universal ancestor was a very different entity than its descendants. If it were a more rudimentary sort of organism, then the tempo of its evolution would have been high and the mode of its evolution highly varied, greatly expanded.

Were the actual root of the universal tree (Fig. 4) located in the vicinity of the deepest branchings in any one of the three primary kingdoms, the above argument concerning sequence distances would not apply to that kingdom, which makes it conceivable that the universal ancestor had the basic phenotype of that group. (This argument is particularly attractive as regards the archaeobacteria, for the group sits relatively close to the intersection of the three primary lineages; see Fig. 4.) However, this would still leave the problem of deriving the other two phenotypes from a third comparably complex one, which entails drastic changes at the molecular level in most functions in the cell. In my opinion the changes in overall cell structure, organization, etc., required to change one of the three phenotypes into either of the others are too drastic and disruptive to have actually occurred.

Accepting all this, the only solution to the problem is for the universal ancestor to have been a progenote. Since the progenote is far simpler and more rudimentary than extant organisms, the significant differences in basic molecular structures and processes that distinguish the three major types of organisms would be attributes that the universal ancestor never possessed. In other words, the more rudimentary versions of a function present in the progenote would become refined and augmented independently, and so uniquely, in each of its progeny lineages. This independent refinement (and augmentation) of a more rudimentary function, not the replacement of one complex function by a different complex version thereof (the beginning stages of which would be strongly selected against), is why remarkable differences in detail have evolved for the basic functions in each of the three kingdoms. Biological specificity does not arise full-blown in cells, and in the transition from the universal ancestor to its descendants we are witnessing the evolution of biological specificity itself.

If the universal ancestor were a progenote, a particular pattern (spectrum) of relationships would exist among the various functions in the three primary kingdoms that would be hard to explain otherwise. The progenote lacked most of the functions characteristic of cells today, and those it did possess existed in a primitive, imprecise form. Therefore, functions that were central to the progenote and its descendants would have undergone the least evolutionary change and so would be the most similar in organisms today. The translation apparatus is a case in point. Without it no genotype-phenotype relationship exists, enzymes as we know them cannot evolve, accurate replication of nucleic acids is impossible, etc. (251). Translation had therefore to

be one of the earliest cellular functions to arise. As then expected, it is one of the most structurally conserved functions. Only the fine-tuning aspects of the process appear to differ from one kingdom to the next; the replacement of the thymidine residue found in eubacterial and eucaryotic tRNAs by 1-methyl-pseudouridine in archaeobacterial tRNAs (73, 165) exemplifies the type of differences encountered.

Structure developed only in crude, primitive ways in the progenote would undergo significant refinement and augmentation in the descendant lineages. While such functions would be homologous in all kingdoms, they would be noticeable idiosyncratic and characteristic in each as well. RNA polymerase could be an example here (281, 282). Functions not present in the progenote would then be totally idiosyncratic or be analogous, not homologous, in kingdom comparisons. Aspects of genome organization may turn out to be examples here, for the progenote did not face the problem of organizing thousands of genes, i.e., of developing an ordered genome structure. It is also possible that some primitive function in the progenote becomes reworked into another function(s) in one (or more) of the primary lineages. In this case, we might expect to find examples of structural homology without functional homology between kingdoms. Some control mechanisms may fall into this class (H. Hartman, personal communication).

The hierarchy of diversification suggested by the progenote, from highly homologous structures, to slightly homologous ones, to analogous ones, to idiosyncratic structures, should define the order in which the various processes arose (or became functionally readapted) during cellular evolution. On the consistency of this picture will turn the validity and utility of the progenote concept.

The progenote is today the end of an evolutionary trail that starts with fact, progresses through inference, and fades into fancy. However, in science endings tend to be beginnings. Within a decade we will have before us at least an order of magnitude more evolutionary information than we now possess and will be able to infer a great deal more with a great deal more assurance than we now can. The root of the universal tree will probably have been determined, many gene families will have been defined, the evolution of genomic organization and of control mechanisms will have become serious problems, the enzymatic capacities of RNA will be more thoroughly elucidated, and the relationship between the evolution of the planet and the life thereon will be much better understood. The concepts of the progenote and of nucleic acid life will have come into their own.

#### ACKNOWLEDGMENTS

I appreciate the help of R. R. Gutell in preparing three of the figures used herein.

My work is supported by grants from the National Aeronautics and Space Administration, the National Science Foundation, and the Office of Naval Research.

#### LITERATURE CITED

1. Achenbach-Richter, L., R. Gupta, K. O. Stetter, and C. R. Woese. 1987. Were the original eubacteria thermophiles? *Syst. Appl. Microbiol.* 9:34-39.
2. Altamura, S., P. Cammarano, and P. Londei. 1986. Archaeobacterial and eukaryotic ribosomal subunits can form active hybrid ribosomes. *FEBS Lett.* 204:129-133.
3. Alvarez, L. 1983. Experimental evidence that an asteroid impact led to the extinction of many species 65 million years ago. *Proc. Natl. Acad. Sci. USA* 80:627-642.
4. Ambler, R. P., M. Daniel, J. Hermoso, T. E. Meyer, R. G. Bartsch, and M. D. Kamen. 1979. Cytochrome  $c_2$  sequence



- variation among the recognised species of purple nonsulphur photosynthetic bacteria. *Nature* (London) 278:659-660.
5. Balch, W. E., G. E. Fox, L. J. Magrum, C. R. Woese, and R. S. Wolfe. 1979. Methanogens: reevaluation of a unique biological group. *Microbiol. Rev.* 43:260-296.
  6. Balch, W. E., L. J. Magrum, G. E. Fox, R. S. Wolfe, and C. R. Woese. 1977. An ancient divergence among the bacteria. *J. Mol. Evol.* 9:305-311.
  7. Barbour, A. G., K.-I. Amano, T. Hackstadt, L. Perry, and H. D. Caldwell. 1982. *Chlamydia trachomatis* has penicillin-binding proteins but not detectable muramic acid. *J. Bacteriol.* 151:420-428.
  8. Benecke, W. 1912. *Bau und Leben der Bakterien*. B. G. Teubner, Leipzig.
  9. Blaylock, B. A., and T. C. Stadtman. 1966. Methane biosynthesis by *Methanosarcina barkeri*: properties of the soluble enzyme system. *Arch. Biochem. Biophys.* 116:138-158.
  10. Bonen, L., and W. F. Doolittle. 1975. On the prokaryotic nature of red algal chloroplasts. *Proc. Natl. Acad. Sci. USA* 72:2310-2314.
  11. Bonen, L., and W. F. Doolittle. 1976. Partial sequences of 16S rRNA and the phylogeny of blue-green algae and chloroplasts. *Nature* (London) 261:669-673.
  12. Bonen, L., and W. F. Doolittle. 1978. Ribosomal RNA homologies and the evolution of the filamentous blue-green bacteria. *J. Mol. Evol.* 10:283-291.
  13. Boone, D. R., S. Worakit, I. M. Mathrani, and R. A. Mah. 1986. Alkaliphilic methanogens from high-pH lake sediments. *Syst. Appl. Microbiol.* 7:230-234.
  14. Boublik, R., M. Wydro, W. Hellmann, and F. Jenkins. 1979. Structure of functional *A. salina*-*E. coli* hybrid ribosome by electron microscopy. *J. Supramol. Struct.* 10:397-404.
  15. Brock, T. D. 1970. *Biology of microorganisms*. Prentice-Hall, Inc., Englewood Cliffs, N.J.
  16. Brock, T. D. 1978. Thermophilic microorganisms and life at high temperatures, p. 174. Springer-Verlag, New York.
  17. Brockmann, H., Jr., and A. Lipinski. 1983. Bacteriochlorophyll g. A new bacteriochlorophyll from *Heliobacterium chlorum*. *Arch. Microbiol.* 136:17-19.
  18. Brooks, B. W., R. G. E. Murray, J. L. Johnson, E. Stackebrandt, C. R. Woese, and G. E. Fox. 1980. Red-pigmented micrococci: a basis for taxonomy. *Int. J. Syst. Bacteriol.* 30:627-646.
  19. Brosius, J., J. L. Palmer, J. P. Kennedy, and H. F. Noller. 1978. Complete nucleotide sequence of a 16S ribosomal RNA gene from *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 75:4801-4805.
  20. Brown, W. M., E. M. Prager, A. Wang, and A. C. Wilson. 1982. Mitochondrial DNA sequences of primates: tempo and mode of evolution. *J. Mol. Evol.* 18:225-239.
  21. Cammarano, P., A. Teichner, and P. Londei. 1986. Intralineage heterogeneity of archaeobacterial ribosomes: evidence for two physicochemically distinct ribosome classes within the third eukaryotic kingdom. *Syst. Appl. Microbiol.* 7:137.
  22. Cammarano, P., A. Teichner, P. Londei, M. Acca, B. Nicolaus, J. L. Sanz, and R. Amls. 1985. Insensitivity of archaeobacterial ribosomes to protein synthesis inhibitors: evolutionary implications. *EMBO J.* 4:811-816.
  23. Campbell, A. 1981. Evolutionary significance of accessory DNA elements in bacteria. *Annu. Rev. Microbiol.* 35:44-83.
  24. Canale-Parola, E. 1978. Motility and chemotaxis of spirochetes. *Annu. Rev. Microbiol.* 32:69-99.
  25. Chatton, E. 1937. *Titres et travaux scientifiques*. Sete, Sottano, Italy.
  26. Chec, T. R., and B. L. Bass. 1986. Biological catalysis by RNA. *Annu. Rev. Biochem.* 55:599-629.
  27. Clayton, B. J., and R. K. Clayton. 1978. Properties of photochemical reaction centers purified from *Rhodospseudomonas gelatinosa*. *Biochim. Biophys. Acta* 501:470-477.
  28. Collins, M. D., and T. A. Langworthy. 1983. Respiratory quinone composition of some acidophilic bacteria. *Syst. Appl. Microbiol.* 4:295-303.
  29. Cowan, S. T. 1962. The microbial species—a macromyth?, p. 433-455. In G. C. Ainsworth and P. H. A. Sneath (ed.), *Microbial classification*. Twelfth Symposium of the Society for General Microbiology. Cambridge University Press, Cambridge.
  30. Daniels, C. J., R. Gupta, and W. F. Doolittle. 1985. Transcription and excision of a large intron in the tRNA<sup>trp</sup> gene of an archaeobacterium *Halobacterium volcanii*. *J. Biol. Chem.* 260:3132-3134.
  31. Darwin, F. 1887. *The life and letters of Charles Darwin*, vol. 3, p. 18. John Murray, London.
  32. De Ley, J. 1978. Modern methods in bacterial taxonomy; evaluation, application, prospects, p. 347-357. In *Proceedings of the 4th International Conference on Plant Pathogenic Bacteria*, Angers, vol. 1. Gilbert-Clarey, Tours, France.
  33. De Ley, J., and J. De Smedt. 1975. Improvements of the membrane filter method for DNA:rRNA hybridization. *Antonie van Leeuwenhoek J. Microbiol. Serol.* 41:287-307.
  34. de Rosa, M., S. de Rosa, A. Gambacorta, L. Minale, R. H. Thomson, and R. D. Worthington. 1977. Caldariellaquinone, a unique benzo-b-thiopen-4,7 quinone from *Caldariella acidophila*, an extremely thermophilic and acidophilic bacterium. *J. Chem. Soc. Perkin Trans.* 1:653-657.
  35. de Rosa, M., A. Gambacorta, and A. Gliozzi. 1986. Structure, biosynthesis, and physicochemical properties of archaeobacterial lipids. *Microbiol. Rev.* 50:70-80.
  36. De Soete, G. 1983. A least squares algorithm for fitting additive trees to proximity data. *Psychometrika* 48:621-626.
  37. Dickerson, R. E. 1980. Structural conservatism in proteins over three billion years: cytochrome with a touch of collagen, p. 227-249. In R. Srinivasan (ed.), *Biomolecular structure, conformation, function and evolution*, vol. 1. Pergamon Press, Oxford.
  38. Doolittle, R. F., and B. Blömbäck. 1964. Amino acid sequence investigations of fibrinopeptide from various mammals; evolutionary implications. *Nature* (London) 202:147-152.
  39. Doolittle, W. F. 1985. Genome structure in archaeobacteria, p. 545-560. In C. R. Woese and R. S. Wolfe (ed.), *The bacteria*, vol. 8. Archaeobacteria. Academic Press, Inc., New York.
  40. Doolittle, W. F., C. R. Woese, M. L. Sogin, L. Bonen, and D. Stahl. 1975. Sequence studies on 16S ribosomal RNA from a blue-green alga. *J. Mol. Evol.* 4:307-315.
  41. Dron, M., M. Rahire, and J.-D. Rochaix. 1982. Sequence of the chloroplast 16S rRNA gene and its surrounding regions of *Chlamydomonas reinhardtii*. *Nucleic Acids Res.* 10:7609-7620.
  42. Elkmann, B., and R. K. Thauer. 1984. Evidence for the involvement and role of a corrinoid enzyme in methane formation from acetate in *Methanosarcina barkeri*. *Arch. Microbiol.* 142:175-179.
  43. Ernst, W. G. 1983. The early earth and the archaean rock record, p. 41-52. In J. W. Schopf (ed.), *Earth's earliest biosphere*. Princeton University Press, Princeton, N.J.
  44. Errede, B., and M. D. Kamen. 1978. Comparative kinetic studies of cytochromes c in reaction with mitochondrial cytochrome c oxidase and reductase. *Biochemistry* 17:1015-1027.
  45. Feick, R. G., and R. C. Fuller. 1984. Topography of the photosynthetic apparatus of *Chloroflexus aurantiacus*. *Biochemistry* 23:3693-3700.
  46. Felsenstein, J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. *Syst. Zool.* 27:401-410.
  47. Felsenstein, J. 1982. Numerical methods for inferring evolutionary trees. *Q. Rev. Biol.* 57:379-404.
  48. Fitch, W. M. 1971. Toward defining the course of evolution: minimum change for a specified tree topology. *Syst. Zool.* 20:406-416.
  49. Fitch, W. M., and E. Margoliash. 1967. Construction of phylogenetic trees. *Science* 155:279-284.
  50. Fowler, V. J., N. Pfennig, W. Schubert, and E. Stackebrandt. 1984. Towards a phylogeny of phototrophic purple sulfur bacteria: 16S rRNA oligonucleotide cataloguing of eleven species of *Chromatiaceae*. *Arch. Microbiol.* 139:382-387.
  51. Fowler, V. J., F. Widdel, N. Pfennig, C. R. Woese, and E. Stackebrandt. 1986. Phylogenetic relationships of sulfate- and sulfur-reducing eubacteria. *Syst. Appl. Microbiol.* 8:32-41.

52. Fox, G. E. 1985. The structure and evolution of archaeobacterial ribosomal RNA, p. 257-310. In C. R. Woese and R. S. Wolfe (ed.), *The bacteria*, vol. 8. Archaeobacteria. Academic Press, Inc., New York.
53. Fox, G. E., K. R. Luehrsens, and C. R. Woese. 1982. Archaeobacterial 5S ribosomal RNA. *Zentralbl. Bakteriell. Parasitenkd. Infektionskr. Hyg. Abt. 1 Orig. Reihe C* 3:330-345.
54. Fox, G. E., L. J. Magrum, W. E. Balch, R. S. Wolfe, and C. R. Woese. 1977. Classification of methanogenic bacteria by 16S ribosomal RNA characterization. *Proc. Natl. Acad. Sci. USA* 74:4537-4541.
55. Fox, G. E., K. J. Peckman, and C. R. Woese. 1977. Comparative cataloging of 16S ribosomal ribonucleic acid: molecular approach to prokaryotic systematics. *Int. J. Syst. Bacteriol.* 27:44-57.
56. Fox, G. E., E. Stackebrandt, R. B. Hespell, J. Gibson, J. Maniloff, T. A. Dyer, R. S. Wolfe, W. E. Balch, R. Tanner, L. Magrum, L. B. Zablen, R. Blakemore, R. Gupta, L. Bonen, B. J. Lewis, D. A. Stahl, K. R. Luehrsens, K. N. Chen, and C. R. Woese. 1980. The phylogeny of prokaryotes. *Science* 209:457-463.
57. Fuchs, G., and E. Stupperich. 1986. Carbon assimilation pathways in archaeobacteria. *Syst. Appl. Microbiol.* 7:364-369.
58. Fuller, R. C., S. G. Sprague, H. Gest, and R. E. Blankenship. 1985. A unique photosynthetic reaction center from *Helio bacterium chlorum*. *FEBS Lett.* 182:345-349.
59. Garrett, A. J., M. J. Harrison, and G. P. Manire. 1974. A search for the bacterial mucopeptide component muramic acid in *Chlamydia*. *J. Gen. Microbiol.* 80:315-318.
60. Gest, H., and J. L. Favinger. 1983. *Helio bacterium chlorum*, an anoxygenic brownish-green photosynthetic bacterium containing a "new" form of bacteriochlorophyll. *Arch. Microbiol.* 136:11-16.
61. Ghosh, A., J. Das, and J. Maniloff. 1977. Lack of repair of ultraviolet light damage in *Mycoplasma gallisepticum*. *J. Mol. Biol.* 116:337-344.
62. Gibbons, N. E., and R. G. E. Murray. 1978. Proposals concerning the higher taxa of bacteria. *Int. J. Syst. Bacteriol.* 28:1-6.
63. Gibson, J., W. Ludwig, E. Stackebrandt, and C. R. Woese. 1985. The phylogeny of the green photosynthetic bacteria: absence of a close relationship between *Chlorobium* and *Chloroflexus*. *Syst. Appl. Microbiol.* 6:152-156.
64. Gibson, J., E. Stackebrandt, L. B. Zablen, R. Gupta, and C. R. Woese. 1979. A phylogenetic analysis of the purple photosynthetic bacteria. *Curr. Microbiol.* 3:59-64.
65. Golding, G. B. 1983. Estimation of DNA and protein sequence divergence: an examination of some assumptions. *Mol. Biol. Evol.* 1:125-142.
66. Goldschmidt, R. 1940. *The material basis of evolution*. Yale University Press, New Haven, Conn.
67. Graf, L., E. Roux, E. Stutz, and H. Kossel. 1982. Nucleotide sequence of a *Euglena gracilis* chloroplast gene coding for the 16S rRNA: homologies to *E. coli* and *Zea mays* chloroplast 16S rRNA. *Nucleic Acids Res.* 10:6369-6381.
68. Green, C. J., G. C. Stewart, M. A. Hollis, B. S. Vold, and K. F. Bott. 1985. Nucleotide sequence of *Bacillus subtilis* ribosomal RNA operon, *rrnB*. *Gene* 37:261-266.
69. Gullik, A., V. Luzzati, M. de Rosa, and A. Gambacorta. 1986. Structure and polymorphism of tetraether lipids from *Sulfolobus solfataricus*. I. Crystallographic analysis. *Syst. Appl. Microbiol.* 7:258.
70. Gupta, R. 1984. *Halobacterium volcanii* tRNAs. *J. Biol. Chem.* 259:9461-9471.
71. Gupta, R. 1985. Transfer ribonucleic acids of archaeobacteria, p. 311-344. In C. R. Woese and R. S. Wolfe (ed.), *The bacteria*, vol. 8. Archaeobacteria. Academic Press, Inc., New York.
72. Gupta, R., A. Broccoll, and C. R. Woese. 1983. Sequence of the 16S ribosomal RNA from *Halobacterium volcanii*, an archaeobacterium. *Science* 221:656-659.
73. Gupta, R., and C. R. Woese. 1980. Unusual modification patterns in the transfer ribonucleic acids of archaeobacteria. *Curr. Microbiol.* 4:245-249.
74. Gutell, R. R., H. F. Noller, and C. R. Woese. 1986. Higher order structure in ribosomal RNA. *EMBO J.* 5:1111-1113.
75. Gutell, R. R., B. Weiser, C. R. Woese, and H. F. Noller. 1985. Comparative anatomy of 16S-like ribosomal RNA. *Prog. Nucleic Acid Res. Mol. Biol.* 32:155-216.
76. Haeckel, E. 1866. *Generelle Morphologie der Organismen*. Verlag Georg Reimer, Berlin.
77. Haldane, J. B. S. 1929. The origin of life. *Ration. Annl.* 1929:148-169.
78. Hallberg, C., and H. Baltcheffsky. 1981. Solubilization and separation of two b-type cytochromes from a carotenoid mutant of *Halobacterium halobium*. *FEBS Lett.* 125:201-204.
79. Hartmann, R., H.-D. Sickinger, and D. Oesterbelt. 1980. Anaerobic growth of halobacteria. *Proc. Natl. Acad. Sci. USA* 77:3821-3825.
80. Hase, T., W. Wakabayashi, H. Matsubara, L. Kercher, D. Oesterbelt, K. K. Rao, and D. O. Hall. 1977. *Halobacterium halobium* ferredoxin: a homologous protein to chloroplast-type ferredoxins. *FEBS Lett.* 77:308-310.
81. Henderson, E., M. Oakes, M. W. Clark, J. A. Lake, A. T. Matheson, and W. Zillig. 1984. A new ribosome structure. *Science* 225:510-512.
82. Hensel, R., W. Demharter, O. Kandler, R. M. Kroppenstedt, and E. Stackebrandt. 1986. Chemotaxonomic and molecular-genetic studies of the genus *Thermus*: evidence for a phylogenetic relationship of *Thermus aquaticus* and *Thermus ruber* to the genus *Deinococcus*. *Int. J. Syst. Bacteriol.* 36:444-453.
83. Hespell, R. B., B. J. Paster, T. J. Macke, and C. R. Woese. 1984. The origin and phylogeny of the bdellovibrios. *Syst. Appl. Microbiol.* 5:196-203.
84. Herzog, M., and L. Marteaux. 1986. Dinoflagellate 17S rRNA sequence inferred from the gene sequence: evolutionary implications. *Proc. Natl. Acad. Sci. USA* 83:8644-8648.
85. Holland, H. D. 1984. *Evolution of the atmosphere and oceans*, p. 12-16. Princeton University Press, Princeton, N.J.
86. Holländer, R. 1978. The cytochromes of *Thermoplasma acidophilum*. *J. Gen. Microbiol.* 108:165-167.
87. Hori, H., and S. Osawa. 1979. Evolutionary change in 5S RNA secondary structure and a phylogenetic tree of 54 5S rRNA species. *Proc. Natl. Acad. Sci. USA* 76:381-385.
88. Horowitz, N. H. 1945. On the evolution of biochemical syntheses. *Proc. Natl. Acad. Sci. USA* 31:153-157.
89. Huber, R., T. A. Langworthy, H. König, M. Thomm, C. R. Woese, U. B. Sleytr, and K. O. Stetter. 1986. *Thermotoga maritima* sp. nov. represents a new genus of unique extremely thermophilic eubacteria growing up to 90°C. *Arch. Microbiol.* 144:324-333.
90. Huet, J., R. Schnabel, A. Sentenac, and W. Zillig. 1983. Archaeobacteria and eukaryotes possess DNA-dependent RNA polymerases of a common type. *EMBO J.* 2:1291-1294.
91. Hul, J., and P. P. Dennis. 1985. Characterization of the ribosomal RNA gene cluster in *Halobacterium cutirubrum*. *J. Biol. Chem.* 260:529-533.
92. Imhoff, J., H. G. Truper, and N. Pfennig. 1984. Rearrangement of the species and genera of the phototrophic "purple nonsulfur bacteria." *Int. J. Syst. Bacteriol.* 34:340-343.
93. Iwami, M., A. Muto, F. Yamao, and S. Osawa. 1984. Nucleotide sequence of the *rrnB* 16S ribosomal RNA gene from *Mycoplasma capricolum*. *Mol. Gen. Genet.* 196:317-322.
94. Jackson, T. J., R. F. Ramaley, and W. G. Meinschein. 1973. *Thermomicrobium*, a new genus of extremely thermophilic bacteria. *Int. J. Syst. Bacteriol.* 23:28-36.
95. Jarrell, K. F., G. D. Sprott, and A. T. Matheson. 1984. Intracellular potassium concentration and relative acidity of the ribosomal proteins of methanogenic bacteria. *Can. J. Microbiol.* 30:663-668.
96. Jarsch, M., and A. Böck. 1985. Sequence of the 16S ribosomal RNA gene from *Methanococcus vannielii*. *Syst. Appl. Microbiol.* 6:54-59.
97. Jensen, R. A. 1985. Biochemical pathways in prokaryotes can be traced backward through evolutionary time. *Mol. Biol. Evol.* 2:87-120.

98. Jones, W. J., J. A. Leigh, F. Mayer, C. R. Woese, and R. S. Wolfe. 1983. *Methanococcus jannaschii* sp. nov., an extremely thermophilic methanogen from a submarine hydrothermal vent. Arch. Microbiol. 136:254-261.
99. Jukes, T. H., and C. R. Cantor. 1969. Evolution of protein molecules, p. 21-132. In H. N. Munro (ed.), Mammalian protein metabolism. Academic Press, Inc., New York.
100. Kalne, B. P., R. Gupta, and C. R. Woese. 1983. Putative introns in tRNA genes of procaryotes. Proc. Natl. Acad. Sci. USA 80:3309-3312.
101. Kandler, O., and H. König. 1985. Cell envelopes of archaeobacteria, p. 413-458. In C. R. Woese and R. S. Wolfe (ed.), The bacteria, vol. 8. Archaeobacteria. Academic Press, Inc., New York.
102. Kandler, O., and W. Zillig (ed.). 1986. Proceedings of the EMBO Workshop on Molecular Genetics of Archaeobacteria and the International Workshop on Biology and Biochemistry of Archaeobacteria, Munich, June 1985. Gustav Fischer Verlag, Stuttgart, Federal Republic of Germany.
103. Kates, M. 1978. The phytanyl ether-linked polar lipids and isoprenoid neutral lipids of extremely halophilic bacteria. Prog. Chem. Fats Other Lipids 15:301-342.
104. Kenyon, D. H., and G. Steinman. 1969. Biochemical predestination. McGraw-Hill Book Co., New York.
105. Kimura, M. 1982. The neutral theory as a basis for understanding the mechanism of evolution and variation at the molecular level, p. 3-58. In M. Kimura (ed.), Molecular evolution, protein polymorphism and the neutral theory. Japan Scientific Societies Press, Tokyo.
106. Kimura, M. 1983. The neutral theory of molecular evolution. Cambridge University Press, New York.
107. Kjems, J., and R. A. Garrett. 1985. An intron in the 23S ribosomal RNA gene of the archaeobacterium *Desulfurococcus mobilis*. Nature (London) 318:675-677.
108. Kluyver, A. J., and C. B. van Niel. 1936. Prospects for a natural system of classification of bacteria. Zentralbl. Bakteriell. Parasitenkd. Infektionskr. Abt. 1 94:369-403.
109. König, H., and K. O. Stetter. 1986. Studies on archaeobacterial S-layers. Syst. Appl. Microbiol. 7:300-309.
110. König, W., H. Schlesner, and P. Hirsch. 1984. Cell wall studies on budding bacteria of the planctomyces-pasteuria group and on a *Prosthecomicrobium* sp. Arch. Microbiol. 138:200-205.
111. Kreisl, P., and O. Kandler. 1986. Chemical structure of the cell wall polymer of *Methanosarcina*. Syst. Appl. Microbiol. 7:293-299.
112. Kuhn, W., K. Fiebig, H. Hippe, R. A. Mah, B. A. Huser, and G. Gottschalk. 1983. Distribution of cytochromes in methanogenic bacteria. FEMS Microbiol. Lett. 20:407-410.
113. Kushner, D. J. 1985. The Halobacteriaceae, p. 171-214. In C. R. Woese and R. S. Wolfe (ed.), The bacteria, vol. 8. Archaeobacteria. Academic Press, Inc., New York.
114. Lake, J. A., M. W. Clark, E. Henderson, S. P. Fay, M. Oakes, A. Schelman, J. P. Thornber, and R. A. Mah. 1985. Eubacteria, halobacteria, and the origin of photosynthesis: the photocycles. Proc. Natl. Acad. Sci. USA 82:3716-3720.
115. Lake, J. A., E. Henderson, M. W. Clark, and M. Oakes. 1984. Eocytes: a new ribosome structure indicates a kingdom with a close relationship to eukaryotes. Proc. Natl. Acad. Sci. USA 81:3786-3790.
116. Lane, D. J., B. Pace, G. J. Olsen, D. A. Stahl, M. L. Sogin, and N. R. Pace. 1985. Rapid determination of 16S ribosomal RNA sequences for phylogenetic analysis. Proc. Natl. Acad. Sci. USA 82:6955-6959.
117. Langworthy, T. A. 1982. Lipids of bacteria living in extreme environments. Curr. Top. Membr. Transp. 17:45-77.
118. Langworthy, T. A. 1985. Lipids of archaeobacteria, p. 459-498. In C. R. Woese and R. S. Wolfe (ed.), The bacteria, vol. 8. Archaeobacteria. Academic Press, Inc., New York.
119. Langworthy, T. A., G. Holzner, J. G. Zeikus, and T. G. Tornabene. 1983. Iso- and anteiso-branched glycerol diethers of the thermophilic anaerobe *Thermodesulfotobacterium commune*. Syst. Appl. Microbiol. 4:1-17.
120. Langworthy, T. A., and J. L. Pond. 1986. Archaeobacterial ether lipids and chemotaxonomy. Syst. Appl. Microbiol. 7:253-257.
121. Lanyi, J. K. 1968. Studies of the electron transport chain of extremely halophilic bacteria. 1. Spectrophotometric identification of the cytochromes of *Halobacterium cutirubrum*. Arch. Biochem. Biophys. 128:716-724.
122. Lanyi, J. K. 1978. Light energy conversion in *Halobacterium halobium*. Microbiol. Rev. 42:682-706.
123. Larsen, H. 1973. The halobacteria's confusion to biology. Antoine van Leeuwenhoek J. Microbiol. Serol. 39:383-396.
124. Lechner, K., G. Wich, and A. Böck. 1985. The nucleotide sequence of the 16S rRNA gene and flanking regions from *Methanobacterium formicicum*: on the phylogenetic relationship between methanogenic and halophilic archaeobacteria. Syst. Appl. Microbiol. 6:157-163.
125. Leffers, H., and R. A. Garrett. 1984. The nucleotide sequence of the 16S ribosomal RNA gene of the archaeobacterium *Halococcus morrhuae*. EMBO J. 3:1613-1619.
126. Leinfelder, W., M. Jarsch, and A. Böck. 1985. The phylogenetic position of the sulfur-dependent archaeobacterium *Thermoproteus tenax*: sequence of the RNA gene. Syst. Appl. Microbiol. 6:164-170.
127. Lewin, R. A. 1981. *Prochloron* and the theory of symbiogenesis. Annu. N.Y. Acad. Sci. 361:325-328.
128. Li, W.-H., C.-C. Luo, and C. Wu. 1985. Evolution of DNA sequences, p. 1-94. In R. J. MacIntyre (ed.), Molecular evolutionary genetics. Plenum Publishing Corp., New York.
129. Li, W.-Y., R. Reddy, D. Henning, P. Epstein, and H. Busch. 1982. Nucleotide sequence of 7S RNA: homology to Alu DNA and LA 4.5S RNA. J. Biol. Chem. 257:5136-5142.
130. Ludwig, W., K.-H. Schleifer, H. Reichenbach, and E. Stackebrandt. 1983. A phylogenetic analysis of the myxobacteria *Myxococcus fulvus*, *Stigmatella aurantiaca*, *Cystobacter fuscus*, *Sorangium cellulosum* and *Nannocystis exedens*. Arch. Microbiol. 135:58-62.
131. Ludwig, W., E. Seewaldt, R. Kilpper-Balz, K.-H. Schleifer, L. Magrum, C. R. Woese, G. E. Fox, and E. Stackebrandt. 1985. The phylogenetic position of *Streptococcus* and *Enterococcus*. J. Gen. Microbiol. 131:543-551.
132. Ludwig, W., and E. Stackebrandt. 1983. A phylogenetic analysis of *Legionella*. Arch. Microbiol. 135:45-50.
133. Luehrsens, K. R., G. E. Fox, M. S. Kilpatrick, R. J. Walker, H. Doomsday, G. Krupp, and H. J. Gross. 1981. The nucleotide sequence of the 5S rRNA from the archaeobacterium *Thermoplasma acidophilum*. Nucleic Acids Res. 9:965-970.
134. Luehrsens, K. R., D. E. Nicholson, Jr., and G. E. Fox. 1985. Widespread distribution of a 7S RNA in archaeobacteria. Curr. Microbiol. 12:69-72.
135. Macario, A. J. L., and E. Conway de Macario. 1985. Antibodies for methanogenic biotechnology. Trends Biotechnol. 3:204-208.
136. Madigan, M. T., and T. D. Brock. 1977. Single "chlorobium-type" vesicles of phototrophically grown *Chloroflexus aurantiacus* observed using negative staining techniques. J. Gen. Microbiol. 102:279-285.
137. Maniloff, J., and H. J. Morowitz. 1972. Cell biology of the mycoplasmas. Bacteriol. Rev. 36:263-290.
138. Margulis, L. 1970. Origin of eucaryotic cells. Yale University Press, New Haven, Conn.
139. Margulis, L., and K. V. Schwartz. 1982. Five kingdoms: an illustrated guide to the phyla of life on earth. W. H. Freeman and Co., New York.
140. Marmur, J., S. Falkow, and M. Mandel. 1963. New approaches to bacterial taxonomy. Annu. Rev. Microbiol. 17:329-372.
141. Matheson, A. T. 1985. Ribosomes of archaeobacteria, p. 345-412. In C. R. Woese and R. S. Wolfe (ed.), The bacteria, vol. 8. Archaeobacteria. Academic Press, Inc., New York.
142. Matheson, A. T., K. A. Louie, and G. N. Henderson. 1986. The evolution of the ribosomal "A" protein (L 12) in archaeobacteria. Syst. Appl. Microbiol. 7:147-150.
143. Mayr, E. 1942. Systematics and the origin of species. Columbia University Press, New York.
144. Mayr, E. 1982. The growth of biological thought: diversity, evolution, and inheritance. Belknap Press, Cambridge, Mass.

145. McCarroll, R., G. J. Olsen, Y. D. Stahl, C. R. Woese, and M. L. Sogin. 1983. Nucleotide sequence of the *Dicryostelium discoideum* small-subunit ribosomal ribonucleic acid inferred from the gene sequence: evolutionary implications. *Biochemistry* 22:5858-5868.
146. McCloskey, J. A. 1986. Nucleoside modification in archaeobacterial transfer RNA. *Syst. Appl. Microbiol.* 7:246-252.
147. Messing, J., J. Carlson, G. Hagen, I. Rubenstein, and A. Oleson. 1984. Cloning and sequencing of the ribosomal RNA genes in maize: the 17S region. *DNA* 3:31-40.
148. Meyer, T. E., M. A. Cusanovich, and M. D. Kamen. 1986. Evidence against use of bacterial amino acid sequence data for construction of all-inclusive phylogenetic trees. *Proc. Natl. Acad. Sci. USA* 83:217-220.
149. Miller, S. L., and L. E. Orgel. 1974. The origin of life on Earth. Prentice-Hall, Inc., Englewood Cliffs, N.J.
150. Mills, L. B., E. J. Stanbridge, W. D. Sedwick, and D. Korn. 1977. Purification and partial characterization of the principle deoxyribonucleic acid polymerase from *Mycoplasma* sp. *J. Bacteriol.* 132:641-649.
151. Moritz, A., and W. Goebel. 1985. Characterization of the 7S RNA and its gene from halobacteria. *Nucleic Acids Res.* 13:6969-6979.
152. Murray, R. G. E. 1974. A place for bacteria in the living world, p. 4-9. In R. E. Buchanan and N. E. Gibbons (ed.), *Bergey's manual of determinative bacteriology*, 8th ed. The Williams & Wilkins Co., Baltimore.
153. Naumann, E., K. Fahlbusch, and G. Gottschalk. 1984. Presence of a trimethylamine:HS-coenzyme M methyltransferase in *Methanobacterium barkeri*. *Arch. Microbiol.* 138:79-83.
154. Noll, K. M., K. L. Rinehart Jr., R. S. Tanner, and R. S. Wolfe. 1986. Structure of component B (7-mercaptoheptanoylthreonine phosphate) of the methylcoenzyme M methylreductase system of *Methanobacterium thermoautotrophicum*. *Proc. Natl. Acad. Sci. USA* 83:4238-4242.
155. Noller, H. F. 1984. Structure of ribosomal RNA. *Annu. Rev. Biochem.* 53:119-162.
156. Nomura, M., S. Mizushima, M. Ozaki, P. Traub, and P. E. Lowry. 1969. Structure and function of ribosomes and their molecular components. Cold Spring Harbor Symp. Quant. Biol. 34:49-61.
157. Olsen, G. J., D. J. Lane, S. J. Giovannoni, and N. R. Pace. 1986. Microbial ecology and evolution: a ribosomal RNA approach. *Annu. Rev. Microbiol.* 40:337-355.
158. Olsen, G. J., N. R. Pace, M. Nuell, B. P. Kalne, R. Gupta, and C. R. Woese. 1985. Sequence of the 16S rRNA gene from the thermoacidophilic archaeobacterium *Sulfolobus solfataricus* and its evolutionary implications. *J. Mol. Evol.* 22:301-307.
159. Oparin, A. I. 1924. *Proiskhozhdenie zhizny*. Moscow. Izd. Moskovskii Rabochii.
160. Oparin, A. I. 1957. The origin of life on Earth, 3rd ed. Oliver and Boyd, Edinburgh.
161. Oren, A., B. J. Paster, and C. R. Woese. 1984. *Haloanaerobiaceae*: a new family of moderately halophilic, obligatory anaerobic bacteria. *Syst. Appl. Microbiol.* 5:71-80.
162. Oyaizu, H., B. Debrunner-Vossbrinck, L. Mandelco, J. A. Studier, and C. R. Woese. 1987. The green non-sulfur bacteria: a deep branching in the eubacterial line of descent. *Syst. Appl. Microbiol.* 9:47-53.
163. Oyaizu, H., and C. R. Woese. 1985. Phylogenetic relationships among the sulfate respiring bacteria, myxobacteria and purple bacteria. *Syst. Appl. Microbiol.* 6:257-263.
164. Pace, N. R., D. A. Stahl, D. J. Lane, and G. J. Olsen. 1986. The analysis of natural microbial populations by ribosomal RNA sequence. *Adv. Microb. Ecol.* 9:1-55.
165. Pang, H., I. Makoto, Y. Kuchino, S. Nishimura, R. Gupta, C. R. Woese, and J. A. McCloskey. 1982. Structure of a modified nucleoside in archaeobacteria tRNA which replaces ribosylthymine. *J. Biol. Chem.* 257:3589-3592.
166. Paster, B. J., W. Ludwig, W. G. Welsburg, E. Stackebrandt, R. B. Hespell, C. M. Hahn, H. Reichenbach, K. O. Stetter, and C. R. Woese. 1985. A phylogenetic grouping of the bacteroides, cytophagae and certain flavobacteria. *Syst. Appl. Microbiol.* 6:34-42.
167. Paster, B. J., E. Stackebrandt, R. B. Hespell, C. M. Hahn, and C. R. Woese. 1985. The phylogeny of spirochetes. *Syst. Appl. Microbiol.* 5:337-351.
168. Pfennig, N. 1977. Phototrophic green and purple bacteria: a comparative, systematic survey. *Annu. Rev. Microbiol.* 31:275-290.
169. Pfennig, N., and H. G. Truper. 1974. The phototrophic bacteria, p. 24-75. In R. E. Buchanan and N. E. Gibbons (ed.), *Bergey's manual of determinative bacteriology*, 8th ed. The Williams & Wilkins Co., Baltimore.
170. Pierson, B. K., and R. W. Castenholz. 1974. A phototrophic gliding filamentous bacterium of hot springs, *Chloroflexus aurantiacus*, gen. and sp. nov. *Arch. Microbiol.* 100:5-24.
171. Pierson, B. K., and J. P. Thornber. 1983. Isolation and spectral characterization of photochemical reaction centers from the thermophilic green bacterium *Chloroflexus aurantiacus* strain J-10-fl. *Proc. Natl. Acad. Sci. USA* 80:80-84.
172. Pond, J. L., T. A. Langworthy, and G. Holtzer. 1986. Long-chain diols: a new class of membrane lipids from a thermophilic bacterium. *Science* 231:1134-1135.
173. Postgate, J. R. 1969. Methane as a minor product of pyruvate metabolism by sulfate-reducing and other bacteria. *J. Gen. Microbiol.* 57:293-302.
174. Qu, L. H., B. Michot, and J.-P. Bachellerie. 1983. Improved methods for structure probing in large RNAs: a rapid "heterologous" sequencing approach is coupled to the direct mapping of nuclease accessible sites. Application to the 5' terminal domain of eukaryotic 28S rRNA. *Nucleic Acids Res.* 11:5903-5919.
175. Raikov, I. B. 1982. The protozoan nucleus: morphology and evolution. *Cell Biol. Monogr.* 9:124-129.
176. Razin, S. 1978. The mycoplasmas. *Microbiol. Rev.* 42:414-470.
177. Reaney, D. C. 1976. Extrachromosomal elements are possible agents of adaptation and development. *Bacteriol. Rev.* 40:552-590.
178. Rogers, M. J., J. Simmons, R. T. Walker, W. G. Welsburg, C. R. Woese, R. S. Tanner, I. M. Robinson, D. A. Stahl, G. Olsen, R. H. Leach, and J. Maniloff. 1985. Construction of the mycoplasma evolutionary tree from 5S rRNA sequence data. *Proc. Natl. Acad. Sci. USA* 82:1160-1164.
179. Rubtsov, P. M., M. M. Musakhanov, V. M. Zakharyev, A. S. Krayev, K. G. Skryabin, and A. A. Bayev. 1980. The structure of the yeast ribosomal RNA genes. I. The complete nucleotide sequence of the 18S ribosomal RNA gene from *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 8:5779-5794.
180. Runnegar, B. 1982. The Cambrian explosion: animals or fossils? *J. Geol. Soc. Aust.* 29:395-411.
181. Salim, M., and E. H. Maden. 1981. Nucleotide sequence of *Xenopus laevis* ribosomal RNA inferred from gene sequence. *Nature (London)* 291:205-208.
182. Sapienza, C., and W. F. Doolittle. 1982. Unusual physical organization of the halobacterial genome. *Nature (London)* 395:384-389.
183. Schleifer, K.-H., and O. Kandler. 1972. Peptidoglycan types of bacterial cell walls and their taxonomic implications. *Bacteriol. Rev.* 36:407-477.
184. Schleifer, K.-H., and E. Stackebrandt. 1983. Molecular systematics of prokaryotes. *Annu. Rev. Microbiol.* 37:143-187.
185. Schleifer, K.-H., J. Steber, and H. Mayer. 1982. Chemical composition and structure of the cell wall of *Halococcus morrhuae*. *Zentralbl. Bakteriol. Parasitenkd. Infektionskr. Hyg. Abt. 1 Orig. Reihe C* 3:171-178.
186. Schlesner, H., and P. Hirsch. 1984. Assignment of ATCC 27377 to *Pirella* gen. nov. as *Pirella staleyii* comb. nov. *Int. J. Syst. Bacteriol.* 34:492-495.
187. Schnabel, R., M. Thomm, R. Gerardy-Schahn, W. Zillig, K. O. Stetter, and J. Huet. 1983. Structural homology between different archaeobacterial DNA-dependent RNA polymerases analyzed by immunological comparison of their components. *EMBO J.* 2:751-755.
188. Schwartz, R. M., and M. O. Dayhoff. 1978. Origins of

- prokaryotes, mitochondria and chloroplasts. *Science* 199:395-403.
189. Schwarz, A., and H. Kossel. 1980. The primary structure of 16S rRNA from *Zea mays* chloroplast is homologous to *E. coli* 16S rRNA. *Nature* (London) 283:739-742.
  190. Searcy, D. G., and F. R. Whatley. 1982. *Thermoplasma acidophilum* cell membrane: cytochrome b and sulfate-stimulated ATPase. *Zentralbl. Bakteri. Parasitenkd. Infektionskr. Hyg. Abt. 1 Orig. Reihe C* 3:245-257.
  191. Seewaldt, E., and E. Stackebrandt. 1982. Partial sequence of 16S ribosomal RNA and the phylogeny of *Prochloron*. *Nature* (London) 295:618-620.
  192. Simpson, G. G. 1944. *Tempo and mode in evolution*. Columbia University Press, New York.
  193. Sneath, P. H. A. 1962. The construction of taxonomic groups, p. 289-332. In G. C. Ainsworth and P. H. A. Sneath (ed.), *Microbial classification. Twelfth Symposium of the Society for General Microbiology*. Cambridge University Press, Cambridge.
  194. Sneath, P. H. A., and R. R. Sokal. 1973. *Numerical taxonomy*. W. H. Freeman and Co., New York.
  195. Sogin, M. L., and H. J. Elwood. 1986. Primary structure of the *Paramecium tetraurelia* small-subunit rRNA coding region: phylogenetic relationships within the ciliophora. *J. Mol. Evol.* 23:53-60.
  196. Sogin, M. L., H. J. Elwood, and J. H. Gunderson. 1986. Evolutionary diversity of eukaryotic small subunit rRNA genes. *Proc. Natl. Acad. Sci. USA* 83:1383-1387.
  197. Sprague, V. 1977. The zoological distribution of microsporidia, p. 335-385. In L. A. Bulla and T. C. Cheng (ed.), *Comparative pathobiology*, vol. 2. Systematics of microsporidia. Plenum Publishing Corp., New York.
  198. Stackebrandt, E., V. J. Fowler, W. Schubert, and J. F. Imhoff. 1984. Towards a phylogeny of phototrophic purple bacteria—the genus *Ectothiorhodospira*. *Arch. Microbiol.* 137:366-370.
  199. Stackebrandt, E., V. J. Fowler, and C. R. Woese. 1983. A phylogenetic analysis of lactobacilli, *Pediococcus pentosaceus* and *Leuconostoc mesenteroides*. *Syst. Appl. Microbiol.* 4:326-337.
  200. Stackebrandt, E., R. M. Kroppenstedt, and V. J. Fowler. 1983. A phylogenetic analysis of the family *Dermatophilaceae*. *J. Gen. Microbiol.* 129:1831-1838.
  201. Stackebrandt, E., B. J. Lewis, and C. R. Woese. 1980. The phylogenetic structure of the coryneform group of bacteria. *Zentralbl. Bakteri. Parasitenkd. Infektionskr. Hyg. Abt. 1 Orig. Reihe C* 101:137-149.
  202. Stackebrandt, E., W. Ludwig, W. Schubert, F. Klink, H. Schlesner, T. Roggentin, and P. Hirsch. 1984. Molecular genetic evidence for early evolutionary origin of budding peptidoglycan-less eubacteria. *Nature* (London) 307:735-737.
  203. Stackebrandt, E., H. Pohl, R. Kroppenstedt, H. Hippe, and C. R. Woese. 1985. 16S rRNA analysis of *Sporomusa*, *Selenomonas*, and *Megasphaera*: on the phylogenetic origin of Gram-positive eubacteria. *Arch. Microbiol.* 143:270-276.
  204. Stackebrandt, E., and C. R. Woese. 1979. A phylogenetic dissection of the family *Micrococcaceae*. *Curr. Microbiol.* 2:317-322.
  205. Stackebrandt, E., and C. R. Woese. 1981. Towards a phylogeny of the Actinomycetes and related organisms. *Curr. Microbiol.* 5:197-202.
  206. Stackebrandt, E., and C. R. Woese. 1981. The evolution of prokaryotes, p. 1-31. In M. J. Carlile, J. R. Collins, and B. E. B. Moseley (ed.), *Molecular and cellular aspects of microbial evolution*. Cambridge University Press, Cambridge.
  207. Stackebrandt, E., B. Wummer-Fuessl, V. Fowler, and K.-H. Schleifer. 1981. DNA homologies and ribosomal RNA similarities among spore-forming members of the order *Actinomycetales*. *Int. J. Syst. Bacteriol.* 31:420-431.
  208. Stahl, D. A., K. R. Luehrs, C. R. Woese, and N. R. Pace. 1981. An unusual 5S RNA from *Sulfolobus acidocaldarius*, and its implications for a general 5S RNA structure. *Nucleic Acids Res.* 9:6129-6137.
  209. Stanier, R. Y. 1970. Some aspects of the biology of cells and their possible evolutionary significance, p. 1-38. In H. P. Charles and B. C. J. G. Knight (ed.), *Organization and control in prokaryotic and eukaryotic cells*. Cambridge University Press, Cambridge.
  210. Stanier, R. Y., and C. B. van Niel. 1941. The main outlines of bacterial classification. *J. Bacteriol.* 42:437-466.
  211. Stanier, R. Y., and C. B. van Niel. 1962. The concept of a bacterium. *Arch. Microbiol.* 42:17-35.
  212. Stetter, K. O. 1982. Ultrathin mycelia-forming organisms from submarine volcanic areas having an optimum growth temperature of 105°C. *Nature* (London) 300:258-260.
  213. Stetter, K. O., and G. Gaag. 1983. Reduction of molecular sulphur by methanogenic bacteria. *Nature* (London) 305:309-311.
  214. Stetter, K. O., M. Thomm, J. Winter, G. Wildgruber, H. Huber, and W. Zillig. 1981. *Methanothermobacter fervidus*, sp. nov., a novel extremely thermophilic methanogen isolated from an Icelandic hot spring. *Zentralbl. Bakteri. Parasitenkd. Infektionskr. Hyg. Abt. 1 Orig. Reihe C* 2:166-178.
  215. Stetter, K. O., and W. Zillig. 1985. *Thermoplasma* and the thermophilic sulfur-dependent archaeobacteria, p. 85-170. In C. R. Woese and R. S. Wolfe (ed.), *The bacteria*, vol. 8. Archaeobacteria. Academic Press, Inc., New York.
  216. Stoekenius, W. 1981. Walsby's square bacterium: fine structure of an orthogonal prokaryote. *J. Bacteriol.* 148:352-360.
  217. Stoekenius, W. 1985. The rhodopsin-like pigments of halobacteria: light-energy and signal transducers in an archaeobacterium. *Trends Biochem. Sci.* 10:483-486.
  218. Stöffler, G., and M. Stöffler-Mellicke. 1986. Electron microscopy of archaeobacterial ribosomes. *Syst. Appl. Microbiol.* 7:123-130.
  219. Stöffler-Mellicke, M., C. Böhme, O. Ströbel, A. Böck, and G. Stöffler. 1985. The structure of ribosomal subunits from *Methanococcus vannielii* as determined by electron microscopy. *Science* 231:1306-1308.
  220. Tanner, R. S., E. Stackebrandt, G. E. Fox, R. Gupta, L. J. Magrum, and C. R. Woese. 1982. A phylogenetic analysis of anaerobic eubacteria capable of synthesizing acetate from carbon dioxide. *Curr. Microbiol.* 7:127-132.
  221. Tanner, R. S., E. Stackebrandt, G. E. Fox, and C. R. Woese. 1981. A phylogenetic analysis of *Acetobacterium woodii*, *Clostridium barkeri*, *Clostridium butyricum*, *Clostridium lituseburense*, *Eubacterium limosum*, and *Eubacterium tenue*. *Curr. Microbiol.* 5:35-38.
  222. Tindall, B. J., H. N. M. Ross, and W. D. Grant. 1984. *Naotrobacterium* gen. nov. and *Naotrococcus* gen. nov., genera of haloalkalophilic archaeobacteria. *Syst. Appl. Microbiol.* 5:41-57.
  223. Tohdoh, N., and M. Sugiyama. 1982. The complete nucleotide sequence of a 16S ribosomal RNA gene from tobacco chloroplasts. *Gene* 17:213-218.
  224. Tomlioka, N., and M. Sugiyama. 1983. The complete nucleotide sequence of a 16S ribosomal RNA gene from a blue-green alga, *Anacystis nidulans*. *Mol. Gen. Genet.* 191:46-50.
  225. Tu, J., D. Prangishvili, H. Huber, G. Wildgruber, W. Zillig, and K. O. Stetter. 1982. Taxonomic relations between archaeobacteria including 6 novel genera examined by cross hybridization of DNAs and 16S rRNAs. *J. Mol. Evol.* 18:109-114.
  226. van Niel, C. B. 1946. The classification and natural relationships of bacteria. *Cold Spring Harbor Symp. Quant. Biol.* 11:285-301.
  - 226a. Vossbrinck, C. R., J. V. Maddox, S. Friedman, B. A. Debrunner-Vossbrinck, and C. R. Woese. 1987. Ribosomal RNA sequence suggests microsporidia are extremely ancient eukaryotes. *Nature* (London) 362:411-414.
  227. Vossbrinck, C. R., and C. R. Woese. 1986. Eukaryotic ribosomes that lack a 5.8S RNA. *Nature* (London) 320:287-288.
  228. Walker, J. C. G., K. Klein, M. Schildowski, J. W. Schopf, D. J. Stevenson, and M. R. Walter. 1983. Environmental evolution of the archaen-early proterozoic earth, p. 260-290. In J. W.

- Schopf (ed.), *Earth's earliest biosphere*. Princeton University Press, Princeton, N.J.
229. Wallace, D. C., and H. J. Morowitz. 1973. Genome size and evolution. *Chromosoma* 40:121-126.
  230. Walsby, A. E. 1980. A square bacterium. *Nature* (London) 283:69-71.
  231. Walter, M. R. 1983. Archaen stromatolites: evidence of the Earth's earliest benthos, p. 187-213. In J. W. Schopf (ed.), *Earth's earliest biosphere*. Princeton University Press, Princeton, N.J.
  232. Waugh, D. S., and N. R. Pace. 1985. Catalysis by RNA. *Bioessays* 4:6-61.
  233. Weisburg, W. G., T. P. Hatch, and C. R. Woese. 1986. Eubacterial origin of chlamydiae. *J. Bacteriol.* 167:570-574.
  234. Weisburg, W. G., Y. Oyaizu, H. Oyaizu, and C. R. Woese. 1985. Natural relationship between bacteroides and flavobacteria. *J. Bacteriol.* 164:230-236.
  235. Weisburg, W. G., C. R. Woese, M. E. Dobson, and E. Weiss. 1985. A common origin of rickettsiae and certain plant pathogens. *Science* 230:556-558.
  236. Whitehead, A. N. *Process and reality*. 1929. The Macmillan Co., Toronto.
  237. Whitehead, A. N. *Modes of thought* 1938. The Macmillan Co., Toronto.
  238. Whitman, W. B. 1985. Methanogenic bacteria, p. 1-84. In C. R. Woese and R. S. Wolfe (ed.), *The bacteria*, vol. 8. Archaeobacteria. Academic Press, Inc., New York.
  239. Whitman, W. B., J. Shieh, S. Sohn, D. S. Caras, and U. Premachandran. 1986. Isolation and characterization of 22 mesophilic methanococci. *Syst. Appl. Microbiol.* 7:235-240.
  240. Whittaker, R. H. 1969. New concepts of kingdoms of organisms. *Science* 163:150-160.
  241. Whittaker, R. H. 1977. Broad classification: the kingdoms and the protozoans, p. 1-34. In J. P. Kreier (ed.), *Parasitic protozoa*, vol. 1. Academic Press, Inc., New York.
  242. Wildgruber, G., M. Thomm, H. König, K. Ober, T. Ricchluto, and K. O. Stetter. 1982. *Methanoplanus limicola*, a plate-shaped methanogen representing a novel family, the *Methanoplanaceae*. *Arch. Microbiol.* 132:31-36.
  243. Wilson, A. C., S. S. Carlson, and T. J. White. 1977. Biochemical evolution. *Annu. Rev. Biochem.* 46:473-639.
  244. Woese, C. R. 1965. On the evolution of the genetic code. *Proc. Natl. Acad. Sci. USA* 54:1546-1552.
  245. Woese, C. R. 1967. The present status of the genetic code. *Prog. Nucleic Acid Res. Mol. Biol.* 7:107-172.
  246. Woese, C. R. 1969. The biological significance of the genetic code. *Prog. Mol. Subcell. Biol.* 1:5-46.
  247. Woese, C. R. 1972. The emergence of genetic organization, p. 301-341. In C. Ponnamperna (ed.), *Exobiology*. North-Holland Publishing Co., Amsterdam.
  248. Woese, C. R. 1973. Evolution of the genetic code. *Naturwissenschaften* 60:447-459.
  249. Woese, C. R. 1979. A proposal concerning the origin of life on the planet Earth. *J. Mol. Evol.* 13:95-101.
  250. Woese, C. R. 1982. Archaeobacteria and cellular origins: an overview. *Zentralbl. Bakteriell. Parasitenkd. Infektionskr. Hyg. Abt. 1 Orig. Reihe C* 3:1-17.
  251. Woese, C. R. 1983. The primary lines of descent and the universal ancestor, p. 209-233. In D. S. Bendall (ed.), *Evolution from molecules to men*. Cambridge University Press, Cambridge.
  252. Woese, C. R. 1985. Why study evolutionary relationships among bacteria?, p. 1-30. In E. Stackebrandt and K. O. Schleifer (ed.), *Evolution of prokaryotes*. Academic Press, Inc. (London), Ltd., London.
  253. Woese, C. R., P. Blanz, and C. M. Hahn. 1984. What isn't a pseudomonad: the importance of nomenclature in bacterial classification. *Syst. Appl. Microbiol.* 5:179-195.
  254. Woese, C. R., P. Blanz, R. B. Hespell, and C. M. Hahn. 1982. Phylogenetic relationships among various helical bacteria. *Curr. Microbiol.* 7:119-124.
  255. Woese, C. R., B. Debrunner-Vossbrinck, H. Oyaizu, E. Stackebrandt, and W. Ludwig. 1985. Gram-positive bacteria: possible photosynthetic ancestry. *Science* 229:762-765.
  256. Woese, C. R., and G. E. Fox. 1977. The concept of cellular evolution. *J. Mol. Evol.* 10:1-6.
  257. Woese, C. R., and G. E. Fox. 1977. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc. Natl. Acad. Sci. USA* 74:5088-5090.
  258. Woese, C. R., J. Gibson, and G. E. Fox. 1980. Do genealogical patterns in purple photosynthetic bacteria reflect interspecific gene transfer? *Nature* (London) 283:212-214.
  259. Woese, C. R., R. Gupta, C. M. Hahn, W. Zillig, and J. Tu. 1984. The phylogenetic relationships of three sulfur-dependent archaeobacteria. *Syst. Appl. Microbiol.* 5:97-105.
  260. Woese, C. R., R. Gutell, R. Gupta, and H. F. Noller. 1983. Detailed analysis of the higher-order structure of 16S-like ribosomal ribonucleic acids. *Microbiol. Rev.* 47:621-669.
  261. Woese, C. R., L. J. Magrum, and G. E. Fox. 1978. Archaeobacteria. *J. Mol. Evol.* 11:245-252.
  262. Woese, C. R., J. Maniloff, and L. B. Zablen. 1980. Phylogenetic analysis of the mycoplasmas. *Proc. Natl. Acad. Sci. USA* 77:494-498.
  263. Woese, C. R., and G. J. Olsen. 1986. Archaeobacterial phylogeny: perspectives on the urkingdoms. *Syst. Appl. Microbiol.* 7:161-177.
  264. Woese, C. R., M. L. Sogin, D. A. Stahl, B. J. Lewis, and L. Bonen. 1976. A comparison of the 16S ribosomal RNAs from mesophilic and thermophilic bacilli. *J. Mol. Evol.* 7:197-213.
  265. Woese, C. R., E. Stackebrandt, and W. Ludwig. 1985. What are mycoplasmas: the relationship of tempo and mode in bacterial evolution. *J. Mol. Evol.* 21:305-316.
  266. Woese, C. R., E. Stackebrandt, T. J. Macke, and G. E. Fox. 1985. A phylogenetic definition of the major eubacterial taxa. *Syst. Appl. Microbiol.* 6:143-151.
  267. Woese, C. R., E. Stackebrandt, W. G. Weisburg, B. J. Paster, M. T. Madigan, V. J. Fowler, C. M. Hahn, P. Blanz, R. Gupta, K. H. Nealson, and G. E. Fox. 1984. The phylogeny of purple bacteria: the alpha subdivision. *Syst. Appl. Microbiol.* 5:315-326.
  268. Woese, C. R., W. G. Weisburg, C. M. Hahn, B. J. Paster, L. B. Zablen, B. J. Lewis, T. J. Macke, W. Ludwig, and E. Stackebrandt. 1985. The phylogeny of purple bacteria: the gamma subdivision. *Syst. Appl. Microbiol.* 6:25-33.
  269. Woese, C. R., W. G. Weisburg, B. J. Paster, C. M. Hahn, R. S. Tanner, N. R. Krieg, H.-P. Kooops, H. Harms, and E. Stackebrandt. 1984. The phylogeny of purple bacteria: the beta subdivision. *Syst. Appl. Microbiol.* 5:327-336.
  270. Woese, C. R., and R. S. Wolfe (ed.). 1985. *The bacteria*, vol. 8. Archaeobacteria. Academic Press, Inc., New York.
  271. Worsley, T. R., R. D. Nance, and J. B. Moody. 1986. Tectonic cycles and the history of the Earth's biogeochemical and paleoceanographic record. *Paleoceanography* 1:233-263.
  272. Wright, S. 1982. The shifting balance theory and macroevolution. *Annu. Rev. Genet.* 16:1-16.
  273. Yamazumi, Z., M. Ihara, Y. Kuchino, R. Gupta, C. R. Woese, and S. Nishimura. 1982. Archaeobacterial tRNA contains 1-methylinosine at residue 57 in the T $\psi$ C loop. *Nucleic Acids Res. Symp. Ser.* 11:209-213.
  274. Yang, D., Y. Oyaizu, H. Oyaizu, G. J. Olsen, and C. R. Woese. 1985. Mitochondrial origins. *Proc. Natl. Acad. Sci. USA* 82:4443-4447.
  275. Yang, D. C., B. Kalne, and C. R. Woese. 1985. The phylogeny of archaeobacteria. *Syst. Appl. Microbiol.* 6:251-256.
  276. Zablen, L. B., M. S. Kissel, C. R. Woese, and D. E. Beutow. 1975. Phylogenetic origin of the chloroplast and procaryotic nature of its ribosomal RNA. *Proc. Natl. Acad. Sci. USA* 72:2418-2422.
  277. Zeikus, J. G. 1977. The biology of methanogenic bacteria. *Bacteriol. Rev.* 41:514-541.
  278. Zeikus, J. G., and R. S. Wolfe. 1972. *Methanobacterium thermoautotrophicum* sp. n., an anaerobic, autotrophic, extreme thermophile. *J. Bacteriol.* 109:707-713.
  279. Zillig, W., F. Gropp, A. Henschen, H. Neumann, P. Palm, W.-D. Reiter, M. Rettenberger, H. Schnabel, and S. Yeats. 1986. Archaeobacterial virus host systems. *Syst. Appl. Micro-*

- biol. 7:58-66.
280. Zillig, W., I. Holz, G. Janekovic, W. Schafer, and W. D. Reiter. 1983. The archaeobacterium *Thermococcus celer* represents a novel genus within the thermophilic branch of the archaeobacteria. Syst. Appl. Microbiol. 4:88-94.
281. Zillig, W., R. Schnabel, and K. O. Stetter. 1985. Archaeobacteria and the origin of the eukaryotic cytoplasm. Curr. Top. Microbiol. Immunol. 114:1-18.
282. Zillig, W., R. Schnabel, J. Tu, and K. O. Stetter. 1982. The phylogeny of archaeobacteria, including novel anaerobic thermoacidophiles, in the light of RNA polymerase structure. Naturwissenschaften 69:197-204.
283. Zuckerkandl, E., and L. Pauling. 1965. Molecules as documents of evolutionary history. J. Theor. Biol. 8:357-366.

## Proposal for a New Hierarchic Classification System, *Actinobacteria* classis nov.

ERKO STACKEBRANDT,\* FRED A. RAINEY, AND NAOMI L. WARD-RAINEY

DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, 38124 Braunschweig, Germany

A new hierarchic classification structure for the taxa between the taxonomic levels of genus and class is proposed for the actinomycete line of descent as defined by analysis of small subunit (16S) rRNA and genes coding for this molecule (rDNA). While the traditional circumscription of a genus of the actinomycete subphylum is by and large in accord with the 16S rRNA/rDNA-based phylogenetic clustering of these organisms, most of the higher taxa proposed in the past do not take into account the phylogenetic clustering of genera. The rich chemical, morphological and physiological diversity of phylogenetically closely related genera makes the description of families and higher taxa so broad that they become meaningless for the description of the enclosed taxa. Here we present a classification system in which phylogenetically neighboring taxa at the genus level are clustered into families, suborders, orders, subclasses, and a class irrespective of those phenotypic characteristics on which the delineation of taxa has been based in the past. Rather than being based on a listing of a wide array of chemotaxonomic, morphological, and physiological properties, the delineation is based solely on 16S rDNA/rRNA sequence-based phylogenetic clustering and the presence of taxon-specific 16S rDNA/RNA signature nucleotides.

In their publication "On the nature of global classification," Wheelis et al. (177) based the definition of higher taxa on the molecular level of universally homologous functions. This statement is derived from the high correlation of genealogical trees inferred from several such molecules, e.g., genes coding for 16S rRNA (16S rDNA) (179), 23S rDNA (96), elongation factors involved in the translation process, and the  $\beta$ -subunit of ATPase (97). The authors (177) stress that a basic requirement of a global classification is uniformity in methods and characteristics used in defining and ranking taxa. Nonhomologous characteristics, on the other hand, may be useful in confirming the molecular groupings. Application of this classification strategy led to the description of domains for the three highest taxa recognized today, the *Archaea*, *Bacteria*, and *Eucarya* (180). As a consequence of the description of kingdoms for the major lineages within the domain *Eucarya* (plants, animals, fungi, and protozoa), Woese et al. (180) described the two main lineages within the domain *Archaea* as the kingdom *Crenarchaeota* and the kingdom *Euryarchaeota*.

Within the domain *Bacteria*, more than 15 lineages, which in phylogenetic uniqueness and ancestry are comparable to the archaeal kingdoms, have been identified. The taxonomic rank of kingdom has not yet been proposed for any of these lineages. The taxon class *Proteobacteria* has been proposed for a phylogenetically broad cluster of gram-negative genera, and several orders have been described for some of the phylogenetic lineages that emerged from the comparison of evolutionarily conserved macromolecules, e.g., *Aquificales* (15), *Thermotogales* (67), *Verrucomicrobiales* (173), and *Planctomycetales* (138). These phylogenetically coherent taxa are now used side by side with higher taxa that were described at the beginning of the pre-molecular era, i.e., before or around 1984. While the phylogenetic coherence of the division *Firmacutes* (53), the class *Mollicutes*, and the orders *Chlamydiales*, *Spirochaetales*, and *Myxobacterales* were by and large confirmed following 16S

rDNA analyses of their members, the majority of higher taxa represent a collection of phylogenetically diverse families and genera. Examples are the classes *Actinomycetes* (81) and *Photobacteria* (53) and the orders *Clostridiales* and *Bacillales* (119), which need to be redefined in order to make classification consistent with phylogenetic structure.

One of the main lines of descent within the domain *Bacteria* includes a wide range of morphologically diverse organisms, most of which, on the basis of a gram-positive staining reaction, can be considered members of the division *Firmacutes* (53). This lineage comprises organisms with a DNA base composition which generally is above 50 mol% G+C (with a few exceptions) and includes representatives of the class *Actinomycetes* (81), the orders *Actinomycetales* (13) and *Micrococcales* (118), the tribes *Brevibacteriae* and *Micrococceae* (120), and several families of the order *Actinomycetales* as well as additional organisms which were identified as members of this lineage by phylogenetic analyses. This lineage encompasses a wide range of bacteria that irrespective of Gram stain reaction, base composition of DNA, morphology, chemotaxonomic properties, and other characteristics used to delineate bacterial taxa in the past, have a common ancestry (Fig. 1).

The modern era in the classification of organisms that are proposed as members of the class *Actinobacteria* has its origin in three sources: firstly, the establishment of chemotaxonomy that detects differences in the chemical composition of cell constituents such as peptidoglycan, polar lipids and fatty acids, isoprenoid quinones, cytochromes, and the base composition of DNA; secondly, the introduction of DNA-DNA reassociation experiments that measure the gross similarities between single-stranded DNA of strains of closely related species (144); and thirdly, the determination of 16S rRNA and rDNA sequence similarities, which reveals the extent of sequence variation among strains at all levels of relatedness (148). Each of these approaches has contributed to the success of a classification strategy which has been termed polyphasic by Colwell (34).

Although the appropriate methods have been available for decades, it took about 30 years to achieve a comprehensive overview of the relatedness among actinomycete bacteria that would allow a proposal for a unified classification system. The

\* Corresponding author. Mailing address: DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Mascheroder Weg 1b, 38124 Braunschweig, Germany. Phone: 49 531 2616 352. Fax: 49 531 2616 418. E-mail: erko@gbf-braunschweig.de.



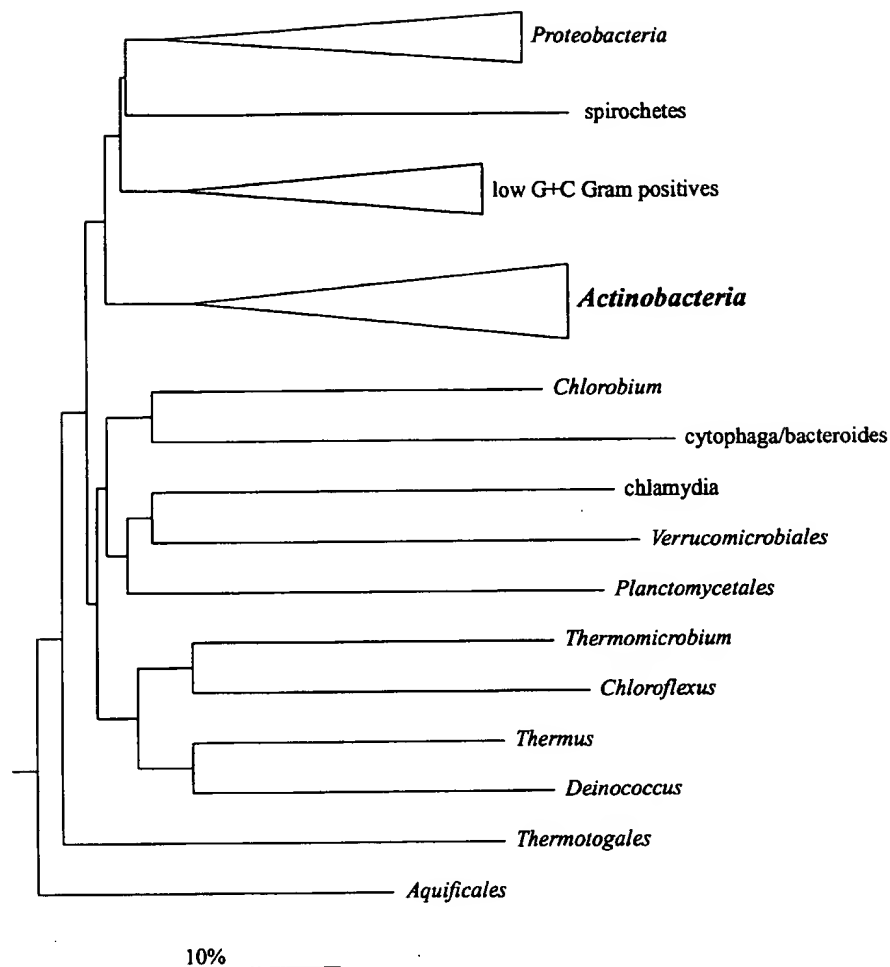


FIG. 1. Dendrogram of 16S rDNA/rRNA relationships showing the phylogenetic position of the proposed class *Actinobacteria* within the domain *Bacteria*. Sequences of *Archaea* were used to root the dendrogram. The recovery of the main lines of descent is as published by Woese (179). The scale bar represents 10 nucleotide substitutions per 100 nucleotides.

reasons were manifold. The most obvious one is the late introduction of 16S rRNA/DNA sequence analysis that, better than any other taxonomic method, places an organism in the framework of phylogenetic relationships. The restricted information on the degree of relatedness obtained by DNA-DNA reassociation studies (143) only allowed analysis of highly related species, while elucidation of more remote relationships had to wait for analysis of conservative genes. Another reason was the tremendous amount of work needed to analyze the phylogenetic relatedness of the type strains of several hundred species which in the end led to the recognition of those properties at the epigenetic level that could be used to unambiguously circumscribe species at the genus level.

#### MATERIALS AND METHODS

**Phylogenetic analyses.** As this publication reviews published data, no emphasis is placed on methods. Rather we refer to relevant publications, most of which contain original or cited data on phylogeny, chemotaxonomy, and other phenotypic data used in the classification of actinomycetes above the species level.

An alignment of all 16S rRNA/DNA sequences currently available for members of the actinomycete lineage was created by using the ac2 editor (99). The sequences included in this alignment were obtained from the Ribosomal Database Project (99) and the ARB package (a software environment for sequence data) (156) as well as our own entries. The 16S rRNA/DNA sequences were

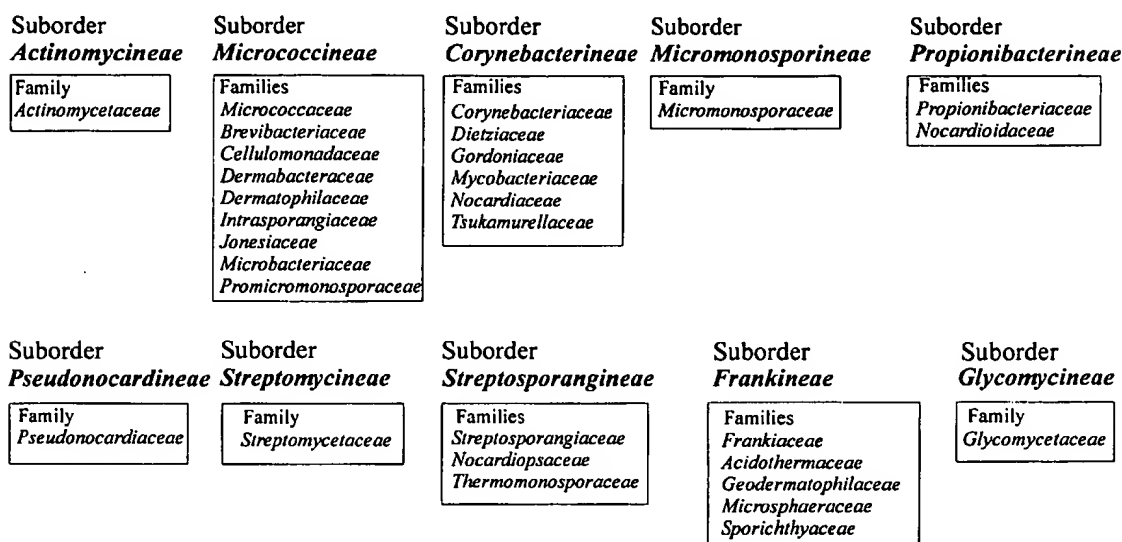
manually aligned to provide a secondary structure-based optimal alignment. The best sequences available were chosen in all cases, but due to the fact that many of the actinomycete reference sequences obtained from the databases are partial sequences consisting of fewer than 1,300 nucleotides, an ideal data set could not be constructed.

The complete data set used for the analyses described in this study contained information on more than 1,000 unambiguous nucleotide positions present in all sequences between positions 150 and 1400 (*Escherichia coli* sequence [11]). Individual data sets, which in many cases comprised more than 1,300 nucleotides, were used for the determination of the relationships within the groups of the proposed taxonomic structure of the actinomycetes. The data set consisting of the type strains of the type species of the genera included in this study can be requested from the authors. The following genera have either not yet been investigated by 16S rDNA sequence analysis and their phylogenetic affiliation remains undetermined or their exact phylogenetic position is unclear: *Actinocorallia* (68), *Actinokinecospira* (64), *Excellispora* (1), *Kineococcus* (186), *Kineospora* (112), and *Micropolyspora* (93).

For the reconstruction of the phylogenetic dendrograms, evolutionary distances were calculated by the method of Jukes and Cantor (70). Phylogenetic dendrograms were reconstructed by using treeing algorithms contained in the PHYLIP package (46) and the ARB package (156). The robustness of tree topologies was evaluated by bootstrap analyses (45) of the neighbor-joining (136) data by performing 1,000 resamplings. The phylogenetic dendrograms presented in many cases represent a general overview at the higher-taxon level and do not show the individual branching order at the genus and species level. These lower-taxon branching orders are susceptible to changes depending on the data set used in the analysis, but the compositions of taxa above the genus level will not be affected.

## Class *Actinobacteria*

Subclass <i>Acidimicrobidae</i>	Order <i>Acidimicrobiales</i>	Family <i>Acidimicrobiaceae</i>
Subclass <i>Rubrobacteridae</i>	Order <i>Rubrobacterales</i>	Family <i>Rubrobacteraceae</i>
Subclass <i>Coriobacteridae</i>	Order <i>Coriobacteriales</i>	Family <i>Coriobacteriaceae</i>
Subclass <i>Sphaerobacteridae</i>	Order <i>Sphaerobacterales</i>	Family <i>Sphaerobacteraceae</i>
Subclass <i>Actinobacteridae</i>	Order <i>Actinomycetales</i>	



## Order *Bifidobacteriales* Family *Bifidobacteriaceae*

FIG. 2. Proposed hierarchic classification system of the class *Actinobacteria* based on the phylogenetic analyses of the 16S rDNA/rRNA sequence data.

### RESULTS AND DISCUSSION

The basis for the proposal of a novel hierarchic structure (Fig. 2) for the phylogenetically coherent group of actinomycete bacteria and relatives is membership of the same phylogenetic group which was formerly described as a 16S rRNA subdivision or subphylum of the gram-positive bacteria. The common ancestry of the actinomycetes proper and the second subdivision of gram-positive bacteria, defined by clostridia, bacilli, and their relatives, has not yet been convincingly demonstrated by 16S rDNA analyses. However, sequence analyses of glutamine synthetase, glutamate dehydrogenase, and the heat shock protein HSP70 reveal a common ancestry of the gram-positive bacteria (59, 60). The phylogenetic coherence of these organisms supports the description of a kingdom for the *Firmacutes* which would contain two or more classes, one of which embraces the actinomycetes proper. If the results of future studies were to unambiguously demonstrate the common ancestry of members of the two subphyla they could be united under the umbrella of a common higher taxon, the kingdom.

Membership of a new strain to the class *Actinobacteria* is indicated by 16S rDNA sequence similarity values above 80% as determined by comparison of almost-complete 16S rDNA sequences of the new strain and the most deeply branching members of the class, such as *Rubrobacter radiotolerans*, *Acidimicrobium ferrooxidans*, or *Coriobacterium glomerans*, and the presence of signature nucleotides. We are aware that the phylogenetic tree, upon which the conclusions outlined below

are based, is a mathematical model of how bacterial evolution occurs. Signature nucleotides are derivatives of the classification process; i.e., signatures are determined for those organisms that are contained within a particular data set. It is also known that a significant increase in species numbers in any of the phylogenetic lineages may lead to a decrease in the number of signatures as the 16S rDNA of more slowly or more rapidly evolving strains may not contain the signature. It is hoped that this proposal will stimulate analyses of other conservative genes from organisms that cluster together by 16S rDNA analyses, so that taxonomic information is provided from additional, independently selected genes and properties and the hierarchic structure proposed here can be tested.

The signatures given below for the higher taxa were chosen for their presence in more than 95% of the members of the respective taxon. The signature pattern for monospecific families must be considered tentative. It should be mentioned that the pattern of signature nucleotides, but not necessarily each individual nucleotide, is indicative of the membership of a taxon to a higher taxon. Bootstrap values (not shown) were determined for the branching points shown in Fig. 3 and were higher than 90% in only a few cases, indicating a lack of statistical significance of the respective branching points. Despite the finding that the majority of branching points are not supported by high bootstrap values, the orders, suborders, and families described previously and below are consistently recoverable from phylogenetic analyses. The lack of high bootstrap

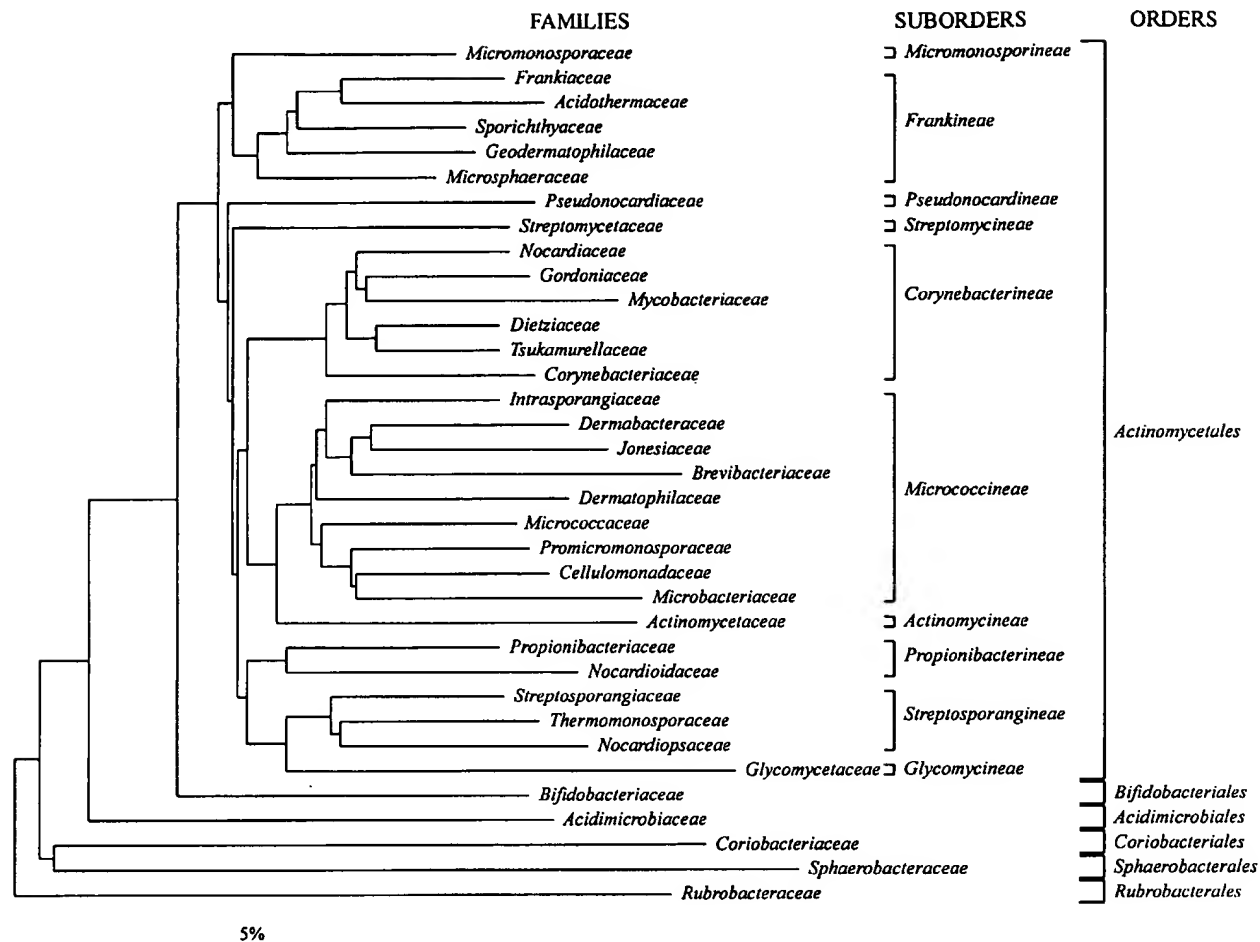


FIG. 3. Intra-class relatedness of *Actinobacteria* showing the presence of six orders as well as the 10 suborders of the order *Actinomycetales* based upon 16S rDNA/rRNA sequence comparison. The phylogenetic relatedness of the families of the class *Actinobacteria* is outlined. The scale bar represents 5 nucleotide substitutions per 100 nucleotides.

values is probably due to the small differences in the primary structure of 16S rDNA of members of most of the neighboring families. Many of the signature nucleotides and idiosyncrasies are located in highly variable and hypervariable regions of the 16S rDNA molecule, which means that dendrograms of relatedness are mostly derived from nucleotide positions that are subject to changes. Also, low bootstrap values are found when the radiation of lineages occurs within a short evolutionary time span, like those found for the radiation of most suborders of the order *Actinomycetales*. Nevertheless, the separate branching of these suborders is found in any of the trees generated, irrespective of the algorithm used. This means that even if the branching points of taxa above the genus level change within the narrow margin at which the higher taxa separate from each other, their phylogenetic uniqueness remains unchanged and their affiliation to the next higher taxon will not be greatly affected. As indicated above, the patterns of signature nucleotides may need to be emended when analysis of novel members of the families indicates the necessity for doing so. In this respect, the taxa defined by molecular data reflect the most recent state-of-the-art insights into molecular systematics. A strategy that could be used in the future to decide whether or not a member of a novel genus could be considered a member of a known or a novel family and/or any

other of the defined higher ranks would include (i) generation of a high-quality, complete or almost-complete 16S rDNA sequence, (ii) proper alignment of the sequence to reference sequences of the same quality, (iii) determination of the phylogenetic position, and (iv) checking for the presence of signature nucleotides as given below for members of the taxon with which the organism groups. If the vast majority of the signature nucleotides match those previously defined for the nearest neighbors, the organism can be considered a member of this taxon. If, however, several signature nucleotides are missing, this genus most likely represents a novel family for which a novel set of signature nucleotides must be defined.

The proposal presented here does not change the current descriptions of species and genera, which are in most cases based upon morphological, chemotaxonomic, and physiological characteristics. These taxa, which provide the working basis for taxonomists, have been revised in the past 20 years to constitute phylogenetically coherent taxonomic units. In contrast, the descriptions of taxa above the genus level have a more broadly based description to incorporate the characteristics of the individual genera contained therein. The description of the family *Micromonosporaceae* should serve as an example (78), and other examples include the families *Propionibacteriaceae* and *Pseudonocardaceae*.

Class *Actinobacteria* classis nov., Stackebrandt, Rainey, and Ward-Rainey. *Actinobacteria* (Ac.ti.no.bac.te'ri.a. Gr. n. *actis*, *actinis*, a ray, beam; Gr. dim. n. *bakterion*, a small rod; *-ia*, proposed ending to denote class; *Actinobacteria*, actinomycete group of bacteria of diverse morphological properties). The class is definable in phylogenetic terms as derived from the analysis of macromolecules of universally homologous functions. Strains of the class *Actinobacteria* can consistently be recovered as members of the same phylogenetic lineage, revealing >80% 16S rDNA/rRNA sequence similarity among each other (Fig. 1), and the presence of the following signature nucleotides in the 16S rDNA/rRNA: an A residue at position 906 and either an A or a C residue at position 955 (except for members of the subclasses *Rubrobacteridae* and *Sphaerobacteridae* [which show U residues at these positions]).

The intraclass relatedness reveals the presence of six phylogenetically distinct lineages which are consistently recovered from phylogenetic analyses (42). These lineages are described as orders (Fig. 3). The 16S rDNA/rRNA signatures defining the higher taxa are based on the available 16S rDNA/rRNA sequences of the type strains of type species of the genera. As certain taxa contain only one or two species, the pattern of signature nucleotides may need to be modified as new species are added to the respective genera.

Subclass *Acidimicrobidae* subclassis nov., Stackebrandt, Rainey, and Ward-Rainey. *Acidimicrobidae* (A.ci.di.mi.cro.bi'dae. M.L. n. *Acidimicrobium*, type genus of the subclass; *-idae*, ending to denote a subclass; M.L. fem. pl. n. *Acidimicrobidae*, the *Acidimicrobium* subclass). The subclass contains the type order *Acidimicrobiales*. The 16S rDNA/rRNA signature pattern is as that of the family *Acidimicrobiaceae*.

Order *Acidimicrobiales* ordo. nov., Stackebrandt, Rainey, and Ward-Rainey. *Acidimicrobiales* (A.ci.di.mi.cro.bi.a'les. M.L. n. *Acidimicrobium*, type genus of the order; *-ales*, ending to denote an order; M.L. fem. pl. n. *Acidimicrobiales*, the *Acidimicrobium* order). The order contains the type family *Acidimicrobiaceae*. The 16S rDNA/rRNA signature pattern is as that of the family *Acidimicrobiaceae*.

Family *Acidimicrobiaceae* fam. nov., Stackebrandt, Rainey, and Ward-Rainey. *Acidimicrobiaceae* (A.ci.di.mi.cro.bi.a'ce.ae. M.L. neut. n. *Acidimicrobium*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Acidimicrobiaceae*, the *Acidimicrobium* family). The following pattern of 16S rDNA/rRNA signature nucleotides and nucleotide pairs defines the family *Acidimicrobiaceae*: 291-309 (U-A), 294-303 (U-A), 408-434 (A-U), 670-736 (C-G), 722-733 (A-C), 955-1225 (C-G), 1118-1155 (C-G), 1311-1326 (A-U), and 1410-1490 (A-U). The family contains the type genus *Acidimicrobium* (24). Phylogenetic analyses have been published previously (17, 23, 149).

Subclass *Rubrobacteridae* subclassis nov., Rainey, Ward-Rainey, and Stackebrandt. *Rubrobacteridae* (Ru.bro.bac.te.ri'dae. M.L. masc. n. *Rubrobacter*, type genus of the subclass; *-idae*, ending to denote a subclass; M.L. fem. pl. n. *Rubrobacteridae*, the *Rubrobacter* subclass). The subclass contains the type order *Rubrobacterales*. The 16S rDNA/rRNA signature pattern is as that of the family *Rubrobacteraceae*.

Order *Rubrobacterales* ordo. nov., Rainey, Ward-Rainey, and Stackebrandt. *Rubrobacterales* (Ru.bro.bac.te.ra'les. M.L. masc. n. *Rubrobacter*, type genus of the order; *-ales*, ending to denote an order; *Rubrobacterales*, the *Rubrobacter* order). The order contains the type family *Rubrobacteraceae*. The 16S rDNA/rRNA signature pattern is as that of the family *Rubrobacteraceae*.

Family *Rubrobacteraceae* fam. nov., Rainey, Ward-Rainey, and Stackebrandt. *Rubrobacteraceae* (Ru.bro.bac.te.ra'ce.ae.

M.L. masc. n. *Rubrobacter*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Rubrobacteraceae*, the *Rubrobacter* family). The following pattern of 16S rDNA/rRNA signature nucleotides and nucleotide pairs defines the family: 127-234 (G-C), 291-309 (U-A), 657-749 (G-C), 681-709 (C-G), 941-1342 (A-U), 955-1225 (U-A), 1051-1207 (C-G), 1115-1185 (C-G), 1311-1326 (A-U), and 1410-1490 (A-U). The family contains the genus *Rubrobacter* (157). A phylogenetic analysis has been published previously (17).

Subclass *Coriobacteridae* subclassis nov., Stackebrandt, Rainey, and Ward-Rainey. *Coriobacteridae* (Co.ri.o.bac.te.ri'dae. M.L. neut. n. *Coriobacterium*, type genus of the subclass; *-idae*, ending to denote a subclass; M.L. fem. pl. n. *Coriobacteridae*, the *Coriobacterium* subclass). The subclass contains the type order *Coriobacterales*. The 16S rDNA/rRNA signature pattern is as that of the family *Coriobacteriaceae*.

Order *Coriobacterales* ordo. nov., Stackebrandt, Rainey, and Ward-Rainey. *Coriobacterales* (Co.ri.o.bac.te.ri.a'les. M.L. neut. n. *Coriobacterium*, type genus of the order; *-ales*, ending to denote an order; M.L. fem. pl. n. *Coriobacterales*, the *Coriobacterium* order). The order contains the type family *Coriobacteriaceae*. The 16S rDNA/rRNA signature pattern is as that of the family *Coriobacteriaceae*.

Family *Coriobacteriaceae* fam. nov., Stackebrandt, Rainey, and Ward-Rainey. *Coriobacteriaceae* (Co.ri.o.bac.te.ri.a'ce.ae. M.L. neut. n. *Coriobacterium*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Coriobacteriaceae*, the *Coriobacterium* family). The pattern of 16S rDNA/rRNA signature nucleotides of members of the family consists of 113-314 (C-G), 294-303 (G-C), 295-302 (U-A), 670-736 (G-C), 771-808 (U-A), 772-807 (A-U), 823-877 (A-U), 941-1342 (A-U), 950-1231 (U-G), 1120-1153 (G-C), 1148 (C), 1165-1171 (C-G), 1242-1295 (G-C), 1313-1324 (G-C), and 1410-1490 (A-U). The family contains the type genus *Coriobacterium* (61) as well as *Atopobium* (26). Phylogenetic analyses have been published previously (128, 146).

Subclass *Sphaerobacteridae* subclassis nov., Stackebrandt, Rainey, and Ward-Rainey. *Sphaerobacteridae* (Sphae.ro.bac.te.ri'dae. M.L. masc. n. *Sphaerobacter*, type genus of the subclass; *-idae*, ending to denote a subclass; M.L. fem. pl. n. *Sphaerobacteridae*, the *Sphaerobacter* subclass). The subclass contains the type order *Sphaerobacterales*. The 16S rDNA/rRNA signature pattern is as that of the family *Sphaerobacteraceae*.

Order *Sphaerobacterales* ordo. nov., Stackebrandt, Rainey, and Ward-Rainey. *Sphaerobacterales* (Sphae.ro.bac.te.ra'les. M.L. masc. n. *Sphaerobacter*, type genus of the order; *-ales*, ending to denote an order; M.L. fem. pl. n. *Sphaerobacterales*, the *Sphaerobacter* order). The order contains the type family *Sphaerobacteraceae*. The 16S rDNA/rRNA signature pattern is as that of the family *Sphaerobacteraceae*.

Family *Sphaerobacteraceae* fam. nov., Stackebrandt, Rainey, and Ward-Rainey. *Sphaerobacteraceae* (Sphae.ro.bac.te.ra'ce.ae. M.L. masc. n. *Sphaerobacter*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Sphaerobacteraceae*, the *Sphaerobacter* family). The 16S rDNA/rRNA signature pattern for the family consists of 291-309 (U-A), 294-303 (U-A), 408-434 (A-U), 417-426 (C-G), 657-749 (G-C), 670-736 (G-C), 681-709 (C-G), 941-1342 (A-U), 955-1225 (U-A), 1120-1153 (G-C), 1148 (C), and 1351-1371 (C-G). The family contains the type genus *Sphaerobacter* (41). A phylogenetic analysis has been published previously (41).

Subclass *Actinobacteridae* subclassis nov., Stackebrandt, Rainey, and Ward-Rainey. *Actinobacteridae* (Ac.ti.no.bac.te.ri'dae. M.L. masc. n. *Actinomyces*, type genus of the subclass; *-idae*, ending to denote a subclass; M.L. fem. pl. n. *Actinobacteridae*, the *Actinomyces* subclass). Members of *Actinobacteri-*

dae can consistently be recovered as members of the same phylogenetic lineage (42, 48, 133) (Fig. 3). The subclass contains two orders, *Actinomycetales* and *Bifidobacteriales*. The type order is *Actinomycetales* Buchanan 1917 (13) emend. Members of the subclass contain an insertion of about 100 bases between helices 54 and 55 within domain III of the 23S rDNA (132). This insertion has not been found in members of the subclass *Coriobacteridae* (cf. reference 42). The pattern of 16S rDNA/rRNA signatures consists of nucleotides at positions 291-309 (C-G), 294-303 (C-G), 408-434 (G-C), 670-736 (A-U), 941-1342 (G-C), 1148 (U), and 1410-1490 (G-C).

**Order Actinomycetales** Buchanan 1917 (13), emend. Stackebrandt, Rainey, and Ward-Rainey. *Actinomycetales* (Ac.ti.no.my.ce.ta'les. M.L. masc. n. *Actinomyces*, type genus of the order; -ales, ending to denote an order; M.L. pl. fem. n. *Actinomycetales*, the *Actinomyces* order). The 16S rDNA signature pattern consists of nucleotides at positions 122-239 (A-G), 449 (A), 450-483 (G-C), 823-877 (G-C), and 1118-1155 (U-A). The order contains the suborders *Actinomycineae*, *Corynebacterineae*, *Frankineae*, *Glycomycineae*, *Micrococcineae*, *Micromonosporineae*, *Propionibacterineae*, *Pseudonocardineae*, *Streptomyceae*, and *Streptosporangineae*. The type suborder is *Actinomycineae*.

**Suborder Actinomycineae** subordo. nov., Stackebrandt, Rainey, and Ward-Rainey. *Actinomycineae* (Ac.ti.no.my.ci'ne.ae. M.L. masc. n. *Actinomyces*, type genus of the suborder; -ineae, ending to denote a suborder; M.L. fem. pl. n. *Actinomycineae*, the *Actinomyces* suborder). The 16S rDNA signature pattern is as that of the type family *Actinomycetaceae*.

**Family Actinomycetaceae** Buchanan 1918 (14), emend. Stackebrandt, Rainey, and Ward-Rainey. *Actinomycetaceae* (Ac.ti.no.my.ce.ta'ce.ae. M.L. masc. n. *Actinomyces*, type genus of the family; -aceae, ending to denote a family; M.L. fem. pl. n. *Actinomycetaceae*, the *Actinomyces* family). The pattern of 16S rDNA signature nucleotides consists of positions 598-640 (U-G), 1059-1198 (U-A), and 1061-1195 (G-U). Genera belonging to the family are the type genus *Actinomyces* (63), *Mobiluncus* (141), and *Arcanobacterium* (32). Phylogenetic analyses have been published previously (51, 87, 99, 142).

**Suborder Propionibacterineae** subordo. nov., Rainey, Ward-Rainey, and Stackebrandt. *Propionibacterineae* (Pro.pi.on.i.bac.te.ri'ne.ae. M.L. neut. n. *Propionibacterium*, type genus of the suborder; -ineae, ending to denote a suborder; M.L. fem. pl. n. *Propionibacterineae*, the *Propionibacterium* suborder). The pattern of 16S rDNA signatures consists of nucleotides at positions 127-234 (A-U), 603-635 (A-U), 657-749 (G-C), 671-735 (A-U), 986-1219 (U-A), 987-1218 (G-C), 990-1215 (U-G), and 1059-1198 (C-G). The suborder contains the type family *Propionibacteriaceae* and the family *Nocardioidaceae*.

**Family Propionibacteriaceae** Delwiche 1957 (40), emend. Rainey, Ward-Rainey, and Stackebrandt. *Propionibacteriaceae* (Pro.pi.o.ni.bac.te.ri.a'ce.ae. M.L. neut. n. *Propionibacterium*, type genus of the family; -aceae, ending to denote a family; M.L. fem. pl. n. *Propionibacteriaceae*, the *Propionibacterium* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 66-103 (A-U), 328 (U), 370-391 (C-G), 407-435 (C-G), 602-636 (A-U), 658-748 (A-U), 686 (G), 780 (A), 787 (C), 819 (G), 825-875 (A-U), and 1409-1491 (A-U). Genera included in the family are the type genus *Propionibacterium* (108) as well as *Luteococcus* (163), *Micrococcus* (103), and *Propionifex* (187). Phylogenetic analyses have been published previously (19, 20, 28, 103, 163, 187).

**Family Nocardioidaceae** Nesterenko et al. 1985 (104), emend. Rainey, Ward-Rainey, and Stackebrandt. *Nocardioidaceae* (No.car.di.o.i.da'ce.ae. M.L. masc. n. *Nocardioides*, type genus of the family; -aceae, ending to denote a family; M.L. fem. pl. n.

*Nocardioidaceae*, the *Nocardioides* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 66-103 (G-C), 328 (C), 370-391 (G-C), 407-435 (A-U), 602-636 (G-U), 658-748 (U-A), 686 (U), 780 (G), 787 (A), 819 (U), 825-875 (G-C), and 1409-1491 (C-G). Genera included in the family are the type genus *Nocardioides* (116) and *Aeromicrobium* (101). Phylogenetic analyses have been published previously (101, 161).

**Suborder Micrococcineae** (*Micrococceae* Prevot 1961) (120), emend. Stackebrandt, Rainey, and Ward-Rainey. *Micrococceae* (Mi.cro.coc.ci'ne.ae. M.L. masc. n. *Micrococcus*, type genus of the suborder; -ineae, ending to denote a suborder; M.L. fem. pl. n. *Micrococcineae*, the *Micrococcus* suborder). The pattern of 16S rDNA signatures consists of nucleotides at positions 66-103 (A-U), 70-98 (U-A), 82-87 (G-C), 127-234 (A-U), 449 (A), 598-640 (U-G), 600-638 (U-G), 722-733 (A-A), 952-1229 (C-G), 986-1219 (A-U), 987-1218 (A-U), and 1059-1198 (U-A). The type family is *Micrococcaceae*. Other families in the suborder include *Cellulomonadaceae*, *Promicromonosporaceae*, *Dermatophilaceae*, *Brevibacteriaceae*, *Dermabacteraceae*, *Intrasporangiaceae*, *Jonesiaceae*, and *Microbacteriaceae*.

**Family Micrococcaceae** Pribham 1929 (121), emend. Stackebrandt, Rainey, and Ward-Rainey. *Micrococcaceae* (Mi.cro.coc.ca'ce.ae. M.L. masc. n. *Micrococcus*, type genus of the family; -aceae, ending to denote a family; M.L. fem. pl. n. *Micrococcaceae*, the *Micrococcus* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 293-304 (G-U), 610 (G), 598-640 (U-U), 615-625 (G-C), 839-847 (A-U), 859 (U), 1025-1036 (C-G), 1026-1035 (C-G), 1265-1270 (U-G), and 1278 (U). The family contains the type genus *Micrococcus* (25) as well as the genera *Arthrobacter* (35; emended in reference 77), *Kocuria* (150), *Nesterenkonia* (150), *Renibacterium* (137), *Rothia* (52), and *Stomatococcus* (7). Phylogenetic analyses have been published previously (76, 150).

**Family Cellulomonadaceae** Stackebrandt and Prauser 1991 (147), emend. Stackebrandt, Rainey, and Ward-Rainey. *Cellulomonadaceae* (Cel.lu.lo.mo.na.da'ce.ae. M.L. fem. n. *Cellulomonas*, type genus of the family; -aceae, ending to denote a family; M.L. fem. pl. n. *Cellulomonadaceae*, the *Cellulomonas* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 30-553 (C-G), 100 (C), 183-194 (A-U), 258-268 (G-C), 610 (A), 615-625 (A-U), 630 (C), 658-748 (G-A), 659-746 (C-G), 694 (G), 747 (C), 832-854 (G-U), 859 (C), 1002-1038 (G-C), 1003-1037 (G-U), 1006-1023 (A-C), and 1256 (C). The family contains the type genus *Cellulomonas* (8; emended in references 24 and 153) as well as the genera *Oerskovia* (117; emended in reference 90) and *Rarobacter* (182). Phylogenetic analyses have been published previously (49, 129, 153).

**Family Promicromonosporaceae** fam. nov., Rainey, Ward-Rainey, and Stackebrandt. *Promicromonosporaceae* (Pro.mi.cro.mo.no.spo.ra'ce.ae. M.L. fem. n. *Promicromonospora*, the type genus of the family; -aceae, ending to denote a family; M.L. fem. pl. n. *Promicromonosporaceae*, the *Promicromonospora* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 77-92 (G-U), 144-178 (U-G), 183-194 (C-G), 199-218 (U-U), 381 (A), 602-636 (G-U), 630 (C), 694 (G), 1002-1038 (Purine-U), 1003-1037 (G-C), 1025-1036 (U-A), and 1267 (U). The family contains the type genus *Promicromonospora* (82). Phylogenetic analyses have been published previously (49, 129).

**Family Dermatophilaceae** Austwick 1958 (4), emend. Stackebrandt, Rainey, and Ward-Rainey. *Dermatophilaceae* (Der.ma.to.phi.la'ce.ae. M.L. masc. n. *Dermatophilus*, type genus of the family; -aceae, ending to denote a family; M.L. fem. pl. n.

*Dermatophilaceae*, the *Dermatophilus* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 146-176 (G-U), 153-168 (G-U), 502-543 (G-C), 546 (G), 580-761 (U-A), 602-636 (C-G), 615-625 (G-C), 659-746 (U-A), 825-875 (G-C), 838-848 (Pyr-Pur), and 1251 (G). The family contains the type genus *Dermatophilus* (170) as well as the genera *Kytococcus* (150) and *Dermacoccus* (10). Phylogenetic analyses have been published previously (150, 151).

**Family *Brevibacteriaceae* Breed 1953 (10), emend. Stackebrandt, Rainey, and Ward-Rainey.** *Brevibacteriaceae* (Brev.i. bac.te.ri.a'ce.ae. M.L. neut. n. *Brevibacterium*, type genus of the family; -aceae, ending to denote a family; M.L. fem. pl. n. *Brevibacteriaceae*, the *Brevibacterium* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 41-401 (U-A), 69-99 (C-U), 142-221 (U-A), 144-178 (U-G), 407-435 (C-G), 586-755 (U-A), 591-648 (G-U), 612-628 (G-C), 616-624 (C-G), 631 (G), 660-745 (A-U), 670-736 (U-A), 896-903 (U-G), 1011-1018 (U-A), 1012-1017 (G-C), 1244-1293 (U-A), 1254-1283 (A-C), 1256 (A), 1257 (G), 1262 (A), 1263-1272 (C-G), 1310-1327 (U-A), and 1442-1460 (U-G). The family contains the type genus *Brevibacterium* (10; emended in reference 31). Phylogenetic analyses have been published previously (16, 127).

**Family *Dermabacteraceae* fam. nov., Stackebrandt, Rainey, and Ward-Rainey.** *Dermabacteraceae* (Der.ma.bac.te.ra'ce.ae. M.L. masc. n. *Dermabacter*, type genus of the family; -aceae, ending to denote a family; M.L. fem. pl. n. *Dermabacteraceae*, the *Dermabacter* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 153-168 (C-G), 248-276 (U-G), 258-268 (A-U), 280 (U), 407-435 (G-U), 580-761 (U-A), 586-755 (U-A), 589-650 (C-G), 602-636 (C-G), 615-625 (A-U), 838-848 (Pur-Pyr), and 1189 (C). The family contains the type genus *Dermabacter* (69) as well as the genus *Brachybacterium* (27). Phylogenetic analyses have been published previously (16, 139).

**Family *Intrasporangiaceae* fam. nov. Rainey, Ward-Rainey, and Stackebrandt.** *Intrasporangiaceae* (In.tra.spo.ran.gi.a'ce. ae. M.L. neut. n. *Intrasporangium*, type genus of the family; -aceae, ending to denote a family; M.L. fem. pl. n. *Intrasporangiaceae*, the *Intrasporangium* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 30-553 (C-G), 69-99 (G-U), 140-223 (G-C), 157-164 (G-C), 258-268 (A-U), 630 (C), 658-748 (G-U), 659-746 (U-A), 660-745 (G-C), 694 (G), 838-848 (C-G), 839-847 (U-A), 859 (C), 1003-1037 (G-C), 1007-1022 (C-G), 1133-1141 (A-U), and 1134-1140 (C-G). The family contains the type genus *Intrasporangium* (71) as well as the genera *Sanguibacter* (47) and *Terrabacter* (29). Phylogenetic analyses have been published previously (29, 47).

**Family *Jonesiaceae* fam. nov., Stackebrandt, Rainey, and Ward-Rainey.** *Jonesiaceae* (Jone.si.a'ce.ae. M.L. fem. n. *Jonesia*, type genus of the family; -aceae, ending to denote a family; M.L. fem. pl. n. *Jonesiaceae*, the *Jonesia* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 153-168 (C-G), 280 (U), 379-384 (G-C), 407-435 (G-U), 445-489 (A-U), 589-650 (U-G), 602-636 (U-G), 615-625 (A-U), 668-738 (U-A), 838-848 (Pur-Pyr), and 1189 (C). The family contains the type genus *Jonesia* (130). Phylogenetic analyses have been published previously (49, 129).

**Family *Microbacteriaceae* Park et al. 1993 (113), emend. Rainey, Ward-Rainey, and Stackebrandt.** *Microbacteriaceae* (Mi.cro.bac.te.ri.a'ce.ae. M.L. masc. n. *Microbacterium*, type genus of the family; -aceae, ending to denote a family; M.L. fem. pl. n. *Microbacteriaceae*, the *Microbacterium* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 45-396 (U-A), 144-178 (C-G), 258-268 (A-U), 497 (A), 615-625 (A-U), 694 (G), 771-808 (G-C), 839-847 (G-U),

1256 (C), 1310-1327 (A-U), and 1414-1486 (U-A). The family contains the type genus *Microbacterium* (109) as well as the genera *Agrococcus* (58), *Agromyces* (54), *Aureobacterium* (30), *Clavibacter* (39), *Curtobacterium* (181), and *Rathayibacter* (189). Phylogenetic analyses have been published previously (58, 122, 158, 159).

**Suborder *Corynebacterineae* subordo. nov., Stackebrandt, Rainey, and Ward-Rainey.** *Corynebacterineae* (Co.ry.ne.bac.te.ri'ne.ae. M.L. masc. n. *Corynebacterium*, type genus of the suborder; -ineae, ending to denote a suborder; M.L. fem. pl. n. *Corynebacterineae*, the *Corynebacterium* suborder). The pattern of 16S rDNA signatures consists of nucleotides at positions 127-234 (G-C), 131-231 (U-Pur), 502-543 (A-U), 658-748 (A-A), 564 (C), 600-638 (G-C), 601-637 (U-G), 660-745 (U-A), 671-735 (C-G), 819 (G), 952-1229 (U-A), 986-1219 (U-A), 1116-1184 (U-G), and 1414-1486 (U-G). The suborder contains the type family *Corynebacteriaceae* as well as the families *Dietziaceae*, *Gordoniaceae*, *Mycobacteriaceae*, *Nocardiaceae*, and *Tsukamurellaceae*.

**Family *Corynebacteriaceae* Lehmann and Neumann 1907 (95), emend. Stackebrandt, Rainey, and Ward-Rainey.** *Corynebacteriaceae* (Co.ry.ne.bac.te.ri.a'ce.ae. M.L. neut. n. *Corynebacterium*, type genus of the family; -aceae, ending to denote a family; M.L. fem. pl. n. *Corynebacteriaceae*, the *Corynebacterium* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 293-304 (G-U), 307 (A), 316-337 (U-G), 468 (U), 508 (U), 586-755 (U-G), 631 (G), 661-744 (G-C), 662-743 (U-G), 771-808 (A-U), 824-876 (C-G), 825-875 (G-C), 837-849 (G-U), 843 (C), and 1059-1198 (U-A). The family contains the type genus *Corynebacterium* (94) as well as the genus *Turicella* (50). Phylogenetic analyses have been published previously (50, 114, 135).

**Family *Mycobacteriaceae* Chester 1897 (21), emend. Stackebrandt, Rainey, and Ward-Rainey.** *Mycobacteriaceae* (My.co.bac.te.ri.a'ce.ae. M.L. neut. n. *Mycobacterium*, type genus of the family; -aceae, ending to denote a family; M.L. fem. pl. n. *Mycobacteriaceae*, the *Mycobacterium* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 70-98 (A-U), 293-304 (G-U), 307 (C), 328 (U), 614-626 (A-U), 631 (G), 661-744 (G-C), 824-876 (U-A), 825-875 (A-U), 843 (C), and 1122-1151 (A-U). The family contains the type genus *Mycobacterium* (94). Phylogenetic analyses have been published previously (115, 131).

**Family *Nocardiaceae* Castellani and Chalmers 1919 (18), emend. Rainey, Ward-Rainey, and Stackebrandt.** *Nocardiaceae* (No.car.di.a'ce.ae. M.L. fem. n. *Nocardia* type genus of the family; -aceae, ending to denote a family; M.L. fem. pl. n. *Nocardiaceae*, the *Nocardia* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 70-98 (U-A), 139-224 (G-C), 843 (C), 1008-1021 (C-G), 1189 (C), 1244-1293 (C-G), and 1308-1329 (C-G). The family contains the type genus *Nocardia* (169) as well as the genus *Rhodococcus* (190). Phylogenetic analyses have been published previously (22, 124, 134).

**Family *Gordoniaceae* fam. nov., Rainey, Ward-Rainey, and Stackebrandt.** *Gordoniaceae* (Gor.do.ni.a'ce.ae. M.L. fem. n. *Gordonia*, type genus of the family; -aceae, ending to denote a family; M.L. fem. pl. n. *Gordoniaceae*, the *Gordonia* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 70-98 (A-U), 293-304 (A-U), 307 (U), 661-744 (A-U), 824-876 (U-A), 825-875 (A-U), 843 (U), 1002-1038 (A-U), 1007-1022 (C-G), 1122-1151 (G-C), and 1124-1149 (A-U). The family contains the type genus *Gordonia* (154). Corrigendum: The name *Gordonia*, not *Gordona*, is proposed as the correct etymology. Phylogenetic analyses have been published previously (6, 124, 134).

**Family Tsukamurellaceae** fam. nov., Rainey, Ward-Rainey, and Stackebrandt. *Tsukamurellaceae* (Tsu.ka.mu.rel.la'ce.ae. M.L. fem. n. *Tsukamurella*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Tsukamurellaceae*, the *Tsukamurella* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 70-98 (U-A), 293-304 (G-U), 307 (C), 631 (C), 661-744 (G-C), 824-876 (U-A), 825-875 (A-U), 843 (C), 1007-1022 (G-U), and 1122-1151 (A-U). The family contains the type genus *Tsukamurella* (33). Phylogenetic analyses have been published previously (33, 124, 183).

**Family Dietziaceae** fam. nov., Rainey, Ward-Rainey, and Stackebrandt. *Dietziaceae* (Diet.zi.a'ce.ae. M.L. fem. n. *Dietzia*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Dietziaceae*, the *Dietzia* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 70-98 (U-A), 293-304 (G-U), 307 (U), 418-425 (U-A), 508 (U), 614-626 (U-G), 631 (G), 661-744 (A-U), 771-808 (A-U), 824-876 (C-G), 825-875 (G-C), 843 (C), 1049-1198 (U-A), and 1122-1151 (A-U). The family contains the type genus *Dietzia* (125). A phylogenetic analysis has been published previously (124).

**Suborder Pseudonocardineae** subordo. nov., Stackebrandt, Rainey, and Ward-Rainey. *Pseudonocardineae* (Pseu.do.no.car'di.ne.ae. M.L. fem. n. *Pseudonocardia*, the type genus of the suborder; *-ineae*, ending to denote a suborder; M.L. fem. pl. n. *Pseudonocardineae*, the *Pseudonocardia* suborder). The pattern of 16S rDNA signatures is as that for the family. The type family is *Pseudonocardiaceae*.

**Family Pseudonocardiaceae** Warwick et al. 1994 (175), emend. Stackebrandt, Rainey, and Ward-Rainey. *Pseudonocardiaceae* (Pseu.do.no.car.di.a'ce.ae. M.L. fem. n. *Pseudonocardia*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Pseudonocardiaceae*, the *Pseudonocardia* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 127-234 (G-C), 183-194 (G-U), 502-543 (A-U), 603-635 (C-G), 610 (A), 747 (A), 952-1229 (U-A), 986-1219 (U-A), 987-1218 (G-C), 1001-1039 (Pyr-G), and 1308-1329 (C-G). Comment: Although in phylogenetic terms this family is rather broad, it is currently not possible to subdivide the family due to the lack of an unambiguous pattern of signature nucleotides. The family contains the type genus *Pseudonocardia* (66) as well as the genera *Actinopolyspora* (55), *Actinosynnema* (65), *Amycolatopsis* (92), *Kibdelosporangium* (140), *Kutzneria* (152), *Lenizea* (184), *Saccharomonospora* (106), *Saccharopolyspora* (86), *Saccharothrix* (84), *Streptoalloteichus* (168), and *Thermocrispum* (79). Phylogenetic analyses have been published previously (9, 43, 74, 79, 175, 184).

**Suborder Streptomycineae** subordo. nov., Rainey, Ward-Rainey, and Stackebrandt. *Streptomycineae* (Strep.to.my.ci'ne.ae. M.L. masc. n. *Streptomyces*, type genus of the suborder; *-ineae*, ending to denote a suborder; M.L. fem. pl. n. *Streptomycineae*, the *Streptomyces* suborder). The pattern of signature nucleotides of 16S rDNA is as that of the type family *Streptomycetaceae*.

**Family Streptomycetaceae** Waksman and Henrici 1943 (171), emend. Rainey, Ward-Rainey, and Stackebrandt. *Streptomycetaceae* (Strep.to.my.ce.ta'ce.ae. M.L. masc. n. *Streptomyces*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Streptomycetaceae*, the *Streptomyces* family). The family is emended to exclude the genus *Sporichthya*. The pattern of 16S rDNA signature nucleotides consists of 71 (G), 80-89 (G-C), 81-88 (C-G), 82-87 (U-G), 127-234 (G-C), 209 (C), 210 (C), 211 (G), 610 (G), 671-735 (U-A), 819 (G), 837-849 (C-G), 950-1231 (U-G), 955-1225 (C-G), 965 (C), 1254-1283 (A-U), and 1409-1491 (C-G). The family contains the type genus *Streptomyces* (171; emended in references 176 and

178). Phylogenetic analyses have been published previously (73, 155, 160, 176, 178).

**Suborder Streptosporangineae** subordo. nov., Ward-Rainey, Rainey, and Stackebrandt. *Streptosporangineae* (Strep.to.spo.ran.gi'ne.ae. M.L. neut. n. *Streptosporangium*, type genus of the suborder; *-ineae*, ending to denote a suborder; M.L. fem. pl. n. *Streptosporangineae*, the *Streptosporangium* suborder). The pattern of 16S rDNA signatures consists of nucleotides at positions 127-234 (A-U), 657-749 (G-Pyr), and 955-1225 (C-G). The type family is *Streptosporangiaceae*.

**Family Streptosporangiaceae** Goodfellow et al. 1990 (56) (validation list no. 34), emend. Ward-Rainey, Rainey, and Stackebrandt. *Streptosporangiaceae* (Strep.to.spo.ran.gi.a'ce.ae. M.L. neut. n. *Streptosporangium*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Streptosporangiaceae*, the *Streptosporangium* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 440-494 (C-G), 445-489 (G-C), 501-544 (C-G), 502-543 (G-C), no extra base between positions 453 and 479, 586-755 (U-G), 613-627 (Pyr-Pur), 681-709 (U-A), 1116-1184 (U-G), 1137 (U), 1355-1367 (A-U), 1436-1465 (G-C), and 1422-1478 (G-U). The family contains the type genus *Streptosporangium* (37) as well as the genera *Herbidospira* (83), *Microbispora* (105), *Microtetraspora* (167), *Planobispora* (164), and *Planomonospora* (166). Phylogenetic analyses have been published previously (172, 174).

**Family Nocardiothraceae** Rainey et al. 1996 (127), emend. Rainey, Ward-Rainey, and Stackebrandt. *Nocardiothraceae* (No.car.di.op.sa'ce.ae. M.L. fem. n. *Nocardiothrix*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Nocardiothraceae*, the *Nocardiothrix* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 440-494 (U-G), 442-492 (C-G), 445-489 (C-G), four extra bases between positions 453 and 479, 501-544 (G-C), 502-543 (A-U), 586-755 (C-G), 603-635 (U-A), 613-627 (C-G), 658-748 (U-A), 671-735 (C-G), 681-709 (U-A), 1003-1037 (U-G), 1116-1184 (C-G), 1137 (A), 1355-1367 (G-C), 1422-1478 (G-U), and 1435-1466 (A-U). The family contains the type genus *Nocardiothrix* (100). A phylogenetic analysis has been published previously (127).

**Family Thermomonosporaceae** fam. nov., Rainey, Ward-Rainey, and Stackebrandt. *Thermomonosporaceae* (Ther.mo.mo.no.spo.ra'ce.ae. M.L. fem. n. *Thermomonospora*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Thermomonosporaceae*, the *Thermomonospora* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 440-494 (C-G), 442-492 (G-C), four to seven extra bases between position 453 and 479, 501-544 (C-G), 502-543 (G-C), 586-755 (C-G), 613-627 (C-G), 658-748 (C-U), 681-709 (C-G), 1003-1037 (A-G), 1116-1184 (C-G), 1355-1367 (A-U), 1422-1478 (G-C), and 1435-1466 (G-C). The family contains the type genus *Thermomonospora* (66) as well as the genera *Actinomadura* (89) and *Spirillospora* (38). Phylogenetic analyses have been published previously (127, 172).

**Suborder Micromonosporineae** subordo. nov., Stackebrandt, Rainey, and Ward-Rainey. *Micromonosporineae* (Mi.cro.mo.no.spo.ri'ne.ae. M.L. fem. n. *Micromonospora*, type genus of the suborder; *-ineae*, ending to denote a suborder; M.L. fem. pl. n. *Micromonosporineae*, the *Micromonospora* suborder). The pattern of 16S rDNA signature nucleotides is as indicated for the family. The suborder contains the type family *Micromonosporaceae*.

**Family Micromonosporaceae** Krassilnikov 1938 (80), emend. Koch et al. 1996 (78), emend. Stackebrandt, Rainey, and Ward-Rainey. *Micromonosporaceae* (Mi.cro.mo.no.spo.ra'ce.ae. M.L. fem. n. *Micromonospora*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Micromonospora*



ceae, the *Micromonospora* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 66-103 (G-C), 127-234 (A-U), 153-168 (C-G), 502-543 (G-C), 589-650 (C-G), 747 (A), 811 (U), 840-846 (C-G), 952-1229 (C-G), 1116-1184 (C-G), and 1133-1141 (G-C). The family contains the type genus *Micromonospora* (111) as well as the genera *Actinoplanes* (36; emended in reference 145), *Catellatospora* (3), *Couchioplanes* (162), *Catenuloplanes* (185), *Dactylosporangium* (165), and *Pilimelia* (72). Phylogenetic analyses have been published previously (75, 78).

**Suborder Frankineae subordo. nov., Stackebrandt, Rainey, and Ward-Rainey.** *Frankineae* (Frank.i'ne.ae. M.L. fem. n. *Frankia*, type genus of the suborder; -ineae, ending to denote a suborder; M.L. fem. pl. n. *Frankineae*, the *Frankia* suborder). The pattern of 16S rDNA signatures consists of nucleotides at positions 82-87 (C-G), 127-234 (G-C), 141-222 (G-C), 371-390 (G-C), 502-543 (A-U), and 1003-1037 (G-G). The suborder contains the type family *Frankiaceae* as well as the families *Acidothermaceae*, *Microsphaeraceae*, *Geodermatophilaceae*, and *Sporichthyaceae*. Phylogenetic analyses have been published previously (44, 62, 107, 123, 126).

**Family Frankiaceae Becking 1970 (5), emend. Hahn et al. 1989 (62), emend. Normand et al. 1996 (107), emend. Stackebrandt, Rainey, and Ward-Rainey.** *Frankiaceae* (Frank.i.a'ce.ae. M.L. fem. n. *Frankia*, type genus of the family; -aceae, ending to denote a family; M.L. fem. pl. n. *Frankiaceae*, the *Frankia* family). The 16S rDNA signature nucleotide pattern consists of 139-224 (G-C), 148-174 (A-G), 155-166 (U-G), 839-847 (A-G), 987-1218 (G-C), 1059-1198 (C-G), and 1308-1329 (C-G). The family contains the type genus *Frankia* (12). Phylogenetic analyses have been published previously (62, 107).

**Family Geodermatophilaceae Normand et al. 1996 (107), emend. Stackebrandt, Rainey, and Ward-Rainey.** *Geodermatophilaceae* (Ge.o.der.ma.to.phi.la'ce.ae. M.L. masc. n. *Geodermatophilus*, type genus of the family; -aceae, ending to denote a family; M.L. fem. pl. n. *Geodermatophilaceae*, the *Geodermatophilus* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 139-224 (C-G), 157-164 (A-U), 158-163 (A-U), 186-191 (C-G), 263 (G), 293-304 (G-U), 986-1219 (U-A), 987-1218 (A-U), 1059-1198 (U-A), and 1308-1329 (U-A). The family contains the type genus *Geodermatophilus* (98) as well as the genus *Blastococcus* (2). Phylogenetic analyses have been published previously (44, 107).

**Family Microsphaeraceae fam. nov., Rainey, Ward-Rainey, and Stackebrandt.** *Microsphaeraceae* (Mi.cro.sphae.ra'ce.ae. M.L. fem. n. *Microsphaera*, type genus of the family; -aceae, ending to denote a family; M.L. fem. pl. n. *Microsphaeraceae*, the *Microsphaera* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 139-224 (C-G), 157-164 (G-C), 186-191 (C-G), 839-847 (U-A), 987-1218 (A-U), 1059-1198 (U-A), and 1308-1329 (U-A). The family contains the type genus *Microsphaera* (188). A phylogenetic analysis has been published previously (188).

**Family Sporichthyaceae fam. nov., Rainey, Ward-Rainey, and Stackebrandt.** *Sporichthyaceae* (Spo.rich.thy.a'ce.ae. M.L. fem. n. *Sporichthya*, type genus of the family; -aceae, ending to denote a family; M.L. fem. pl. n. *Sporichthyaceae*, the *Sporichthya* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 139-224 (U-A), 186-191 (G-C), 600-638 (C-G), 839-847 (U-A), 987-1218 (A-U), 1059-1198 (U-A), and 1308-1329 (U-A). The family contains the type genus *Sporichthya* (91). A phylogenetic analysis has been published previously (126).

**Family Acidothermaceae fam. nov., Rainey, Ward-Rainey, and Stackebrandt.** *Acidothermaceae* (A.ci.do.ther.ma'ce.ae. M.L. masc. n. *Acidothermus*, type genus of the family; -aceae,

ending to denote a family; M.L. fem. pl. n. *Acidothermaceae*, the *Acidothermus* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 139-224 (C-G), 186-191 (G-C), 839-847 (A-U), 987-1218 (G-C), 1059-1198 (C-G), and 1308-1329 (C-G). The family contains the type genus *Acidothermus* (102). A phylogenetic analysis has been published previously (123).

**Suborder Glycomycineae subordo. nov., Rainey, Ward-Rainey, and Stackebrandt.** *Glycomycineae* (Gly.co.my.ci.ne'ae. M.L. masc. n. *Glycomyces*, type genus of the suborder; -ineae, ending to denote a suborder; M.L. fem. pl. n. *Glycomycineae*, the *Glycomyces* suborder). The pattern of signature nucleotides of 16S rDNA is as that of the type family *Glycomycetaceae*.

**Family Glycomycetaceae fam. nov., Rainey, Ward-Rainey, and Stackebrandt.** *Glycomycetaceae* (Gly.co.my.ci.ta'ce.ae. M.L. masc. n. *Glycomyces*, type genus of the family; -aceae, ending to denote a family; M.L. fem. pl. n. *Glycomycetaceae*, the *Glycomyces* family). The 16S rDNA pattern of 16S rDNA signature nucleotides contains 70-98 (A-U), 127-234 (G-Pyr), 140-223 (A-U), 229 (G), 366 (U), 415 (C), 449 (C), 534 (G), 681-709 (A-U), 825-875 (G-C), 999-1041 (C-G), 1059-1198 (C-G), 1064-1192 (G-G), 1117-1183 (A-U), and 1309-1328 (C-G). The family contains the type genus *Glycomyces* (85). The phylogenetic position is shown in Fig. 3.

**Order Bifidobacteriales ordo. nov., Stackebrandt, Rainey, and Ward-Rainey.** *Bifidobacteriales* (Bi.fi.do.bac.te.ri.a'les. M.L. neut. n. *Bifidobacterium*, type genus of the order; -ales, ending to denote an order; M.L. fem. pl. n. *Bifidobacteriales*, the *Bifidobacterium* order). The type family of the order is *Bifidobacteriaceae*. The 16S rDNA nucleotide signature is as that of the family.

**Family Bifidobacteriaceae fam. nov., Stackebrandt, Rainey, and Ward-Rainey.** *Bifidobacteriaceae* (Bi.fi.do.bac.te.ri.a'ce.ae. M.L. neut. n. *Bifidobacterium*, type genus of the family; -aceae, ending to denote a family; M.L. fem. pl. n. *Bifidobacteriaceae*, the *Bifidobacterium* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 122-239 (G-U), 128-233 (C-G), 450-483 (C-G), 602-636 (G-C), 681-709 (C-G), 688-699 (A-U), 823-877 (A-U), 1118-1155 (C-G), and 1311-1326 (A-U). The family contains the type genus *Bifidobacterium* (110) as well as *Gardnerella* (57). A phylogenetic structure of the family has been published previously (88, 99).

#### ACKNOWLEDGMENT

We are indebted to Hans G. Trüper for his advice in the nomenclatural aspect of this work.

#### REFERENCES

- Agre, N. S., and L. N. Guzeva. 1975. New actinomycetes genus *Excellispora* gen. nov. *Mikrobiologiya* 44:518-523.
- Ahrens, R., and G. Moll. 1970. Ein neues knospendes Bakterium aus der Ostsee. *Arch. Mikrobiol.* 70:243-265.
- Asano, K., and I. Kawamoto. 1986. *Catellatospora*, a new genus of *Actinomycetales*. *Int. J. Syst. Bacteriol.* 36:512-517.
- Austwick, P. K. C. 1958. Cutaneous streptothrichosis mycotic dermatitis and strawberry foot rot and the genus *Dermatophilus*. *Van Saceghem Vet. Rev. Annot.* 4:33-48.
- Becking, J. H. 1970. *Frankiaceae* fam. nov. *Actinomycetales* with one new combination and six new species of the genus *Frankia* Brunchorst 1887, 174. *Int. J. Syst. Bacteriol.* 20:201-220.
- Bendiger, B., F. A. Rainey, R. M. Kroppenstedt, M. Moormann, and S. Klatte. 1995. *Gordona hydrophobica* sp. nov., isolated from biofilters for waste gas treatment. *Int. J. Syst. Bacteriol.* 45:544-548.
- Bergan, T., and M. Kocur. 1982. *Stomatococcus mucilaginosus* gen. nov. sp. nov. ep. rev., a member of the family *Micrococcaeae*. *Int. J. Syst. Bacteriol.* 32:374-377.
- Bergey, D. H., F. C. Harrison, R. S. Breed, B. W. Hammer, and F. M. Huntton. 1923. *Bergey's manual of determinative bacteriology*, 1st ed. The Williams and Wilkins Co., Baltimore, Md.



9. Bowen, T., E. Stackebrandt, M. Dorsch, and T. M. Embley. 1989. Transfer of the genera *Kibdelosporangium*, *Saccharothrix* and *Amycolata* into the family *Pseudonocardiaceae*. J. Gen. Microbiol. 135:2529-2536.
10. Breed, R. S. 1953. The families developed from *Bacteriaceae* Cohn with a description of the family *Brevibacteriaceae*. VI Congr. Int. Microbiol. Roma 1:10-15.
11. Brosius, J., M. L. Palmer, J. P. Kennedy, and H. P. Noller. 1978. Complete nucleotide sequence of a 16S ribosomal RNA gene from *Escherichia coli*. Proc. Natl. Acad. Sci. USA 75:4801-4805.
12. Brunchorst, J. 1886. Über einige Wurzelanschwellungen besonders diejenigen von Alnus und den Elaeagnaceen. Bot. Inst. Tübingen 2:151-177.
13. Buchanan, R. E. 1917. Studies in the nomenclature and classification of the bacteria. II. The primary subdivisions of the schizomycetes. J. Bacteriol. 2:155-164.
14. Buchanan, R. E. 1918. Studies in the nomenclature and classification of the bacteria. VIII. The subgroups and genera of the *Actinomycetales*. J. Bacteriol. 3:403-406.
15. Burggraf, S., G. J. Olsen, K. O. Stetter, and C. R. Woese. 1992. A phylogenetic analysis of *Aquifex pyrophilus*. Syst. Appl. Microbiol. 15:352-356.
16. Cai, J., and M. D. Collins. 1994. Phylogenetic analysis of species of the *meso*-diaminopimelic acid-containing genera *Brevibacterium* and *Dermabacter*. Int. J. Syst. Bacteriol. 44:583-585.
17. Carreto, L., E. Moore, M. F. Nobre, R. Wait, P. W. Riley, R. J. Sharp, and M. S. da Costa. 1996. *Rubrobacter xylanophilus* sp. nov., a new thermophilic species isolated from a thermally polluted effluent. Int. J. Syst. Bacteriol. 46:460-465.
18. Castellani, A., and A. J. Chalmers. 1919. Manual of tropical medicine, 3rd ed. Williams, Wood and Co., New York, N.Y.
19. Charfreitag, O., and E. Stackebrandt. 1989. Intra- and intergeneric relationships of the genus *Propionibacterium*. J. Gen. Microbiol. 135:2065-2070.
20. Charfreitag, O., M. D. Collins, and E. Stackebrandt. 1988. Phylogenetic and chemotaxonomic evidence for the reclassification of *Arachnia propionica* as *Propionibacterium propionicus* comb. nov. Int. J. Syst. Bacteriol. 38:354-357.
21. Chester, F. D. 1897. Report of the mycologist; bacteriological work. Del. Agr. Exp. Stn. Bull. 9:38-145.
22. Chun, J., and M. Goodfellow. 1995. A phylogenetic analysis of the genus *Nocardia* with 16S rRNA gene sequences. Int. J. Syst. Bacteriol. 45:240-245.
23. Clark, D. A., and P. R. Norris. 1996. *Acidimicrobium ferrooxidans* gen. nov., sp. nov. mixed-culture ferrous iron oxidation with *Sulfobacillus* species. Microbiology (Reading) 142:785-790.
24. Clark, F. E. 1952. The generic classification of the soil bacteria. Bull. Bact. Nom. Tax. 2:45-56.
25. Cohn, F. 1872. Untersuchungen über Bakterien. Beitr. Biol. Pflanz. 1:127-224.
26. Collins, M. D., and S. Wallbanks. 1992. Comparative sequence analysis of the 16S rRNA genes of *Lactobacillus minutus*, *Lactobacillus rinae* and *Streptococcus parvulus*. Proposal for the creation of a new genus *Atopobium*. FEMS Microbiol. Lett. 95:235-240.
27. Collins, M. D., J. Brown, and D. Jones. 1988. *Brachybacterium faecium* gen. nov., sp. nov., a coryneform bacterium from poultry deep litter. Int. J. Syst. Bacteriol. 38:45-48.
28. Collins, M. D., S. Cockcroft, and S. Wallbanks. 1994. Phylogenetic analysis of a new LL-diaminopimelic acid-containing coryneform bacterium from herbage, *Nocardioides plantarum* sp. nov. Int. J. Syst. Bacteriol. 44:523-526.
29. Collins, M. D., M. Dorsch, and E. Stackebrandt. 1989. Transfer of *Pimelobacter tumescens* to *Terrabacter* gen. nov. as *Terrabacter tumescens* comb. nov. and of *Pimelobacter jensenii* to *Nocardioides* as *Nocardioides jensenii* comb. nov. Int. J. Syst. Bacteriol. 39:1-6.
30. Collins, M. D., D. Jones, R. M. Keddie, R. M. Kroppenstedt, and K. H. Schleifer. 1983. Classification of some coryneform bacteria in a new genus *Aureobacterium*. Syst. Appl. Microbiol. 4:236-252.
31. Collins, M. D., D. Jones, R. M. Keddie, and P. A. H. Sneath. 1980. Reclassification of *Chromobacterium iodinum* Davis in a redefined genus *Brevibacterium* Breed as *Brevibacterium iodinum* nom. rev.; comb. nov. J. Gen. Microbiol. 120:1-10.
32. Collins, M. D., D. Jones, and G. M. Schofield. 1982. Reclassification of "*Corynebacterium haemolyticum*" MacLean, Liebow and Rosenberg in the genus *Arcanobacterium* gen. nov. as *Arcanobacterium haemolyticum* nom. rev., comb. nov. J. Gen. Microbiol. 128:1279-1281.
33. Collins, M. D., J. Smida, M. Dorsch, and E. Stackebrandt. 1988. *Tsukamurella* gen. nov., harboring *Corynebacterium paurometabolum* and *Rhodococcus aurantiacus*. Int. J. Syst. Bacteriol. 38:385-391.
34. Colwell, R. R. 1970. Polyphasic taxonomy of bacteria, p. 421-436. In H. Iizuka and T. Hasegawa (ed.), Proceedings of the International Conference on Culture Collections. Tokyo University Press, Tokyo, Japan.
35. Conn, H. J., and I. Dimmick. 1947. Soil bacteria similar in morphology to *Mycobacterium* and *Corynebacterium*. J. Bacteriol. 54:291-303.
36. Couch, J. N. 1950. *Actinoplanes* a new genus of the *Actinomycetales*. J. Elisha Mitchell Sci. Soc. 66:87-92.
37. Couch, J. N. 1955. A new genus and family of the *Actinomycetales*, with a revision of the genus *Actinoplanes*. J. Elisha Mitchell Sci. Soc. 71:148-155.
38. Couch, J. N. 1963. Some new genera and species of the *Actinoplanaceae*. J. Elisha Mitchell Sci. Soc. 79:53-70.
39. Davis, M. J., A. G. Gillaspie, Jr., A. K. Vidaver, and R. W. Harris. 1984. *Clavibacter*: a new genus containing some phytopathogenic coryneform bacteria, including *Clavibacter xyli* subsp. *xyli* sp. nov., subsp. nov. and *Clavibacter xyli* subsp. *cynodontis* subsp. nov., pathogens that cause ratoon stunting disease of sugarcane and Bermudagrass stunting disease. Int. J. Syst. Bacteriol. 34:107-117.
40. Delwiche, E. A. 1957. Family XI *Propionibacteriaceae*, p. 569. In R. S. Breed, E. G. D. Murray, and N. R. Smith (ed.), Bergey's manual of determinative bacteriology, 7th ed. The Williams and Wilkins Co., Baltimore, Md.
41. Dembarter, W., R. Hensel, J. Smida, and E. Stackebrandt. 1989. *Sphaerobacter thermophilus* gen. nov. spec. nov. a deeply rooting member of the actinomycetes subdivision isolated from thermophilically treated sewage sludge. Syst. Appl. Microbiol. 11:261-266.
42. Embley, T. M., and E. Stackebrandt. 1994. The molecular phylogeny and systematics of the actinomycetes. Annu. Rev. Microbiol. 48:257-289.
43. Embley, T. M., J. Smida, and E. Stackebrandt. 1988. The phylogeny of mycolate-less wall chemotype IV actinomycetes and description of *Pseudonocardiaceae* family nov. Syst. Appl. Microbiol. 11:44-52.
44. Eppard, M., W. E. Krumbein, C. Koch, E. Rhiel, J. T. Staley, and E. Stackebrandt. 1996. Morphological, physiological, and molecular characterization of actinomycetes isolated from dry soil, rocks, and monument surfaces. Arch. Microbiol. 166:12-22.
45. Felsenstein, J. 1985. Confidence limits on phylogenies. An approach using the bootstrap. Evolution 39:783-789.
46. Felsenstein, J. 1993. PHYLIP phylogenetic inference package version 3.5.1. Department of Genetics, University of Washington, Seattle.
47. Fernandez-Garayzabal, J. F., L. Dominguez, C. Pascual, D. Jones, and M. D. Collins. 1995. Phenotypic and phylogenetic characterization of some unknown coryneform bacteria isolated from bovine blood and milk. Description of *Sanguibacter* gen. nov. Lett. Appl. Microbiol. 20:69-75.
48. Fox, G. E., and E. Stackebrandt. 1987. The application of 16S rRNA cataloguing and 5S rRNA sequencing in microbial systematics. Methods Microbiol. 19:405-458.
49. Funke, G., C. P. Ramos, and M. D. Collins. 1995. Identification of some clinical strains of CDC coryneform group A-3 and A-4 bacteria as *Cellulomonas* species and proposal of *Cellulomonas hominis* sp. nov. for some group A-3 strains. J. Clin. Microbiol. 33:2091-2097.
50. Funke, G., S. Stubbs, M. Altwegg, A. Carlotti, and M. D. Collins. 1994. *Turicella otitidis* gen. nov., sp. nov., a coryneform bacterium isolated from patients with otitis media. Int. J. Syst. Bacteriol. 44:270-273.
51. Funke, G., S. Stubbs, A. von Graevenitz, and M. D. Collins. 1994. Assignment of human-derived CDC group 1 coryneform bacteria and CDC group 1-like coryneform bacteria to the genus *Actinomyces* as *Actinomyces neuii* subsp. *neuii* sp. nov., subsp. nov., and *Actinomyces neuii* subsp. *anitratus* subsp. nov. Int. J. Syst. Bacteriol. 44:167-171.
52. Georg, L. K., and J. M. Brown. 1967. *Rothia*, gen. nov. An aerobic genus of the family *Actinomycetaceae*. Int. J. Syst. Bacteriol. 17:79-88.
53. Gibbons, N. E., and R. G. E. Murray. 1978. Proposals concerning the higher taxa of bacteria. Int. J. Syst. Bacteriol. 28:1-6.
54. Gledhill, W. E., and L. E. Casida, Jr. 1969. Predominant catalase-negative soil bacteria. III. *Agronomyces*, gen. n., microorganisms intermediary to *Actinomyces* and *Nocardia*. Appl. Microbiol. 18:340-349.
55. Gochnauer, M. B., G. G. Leppard, P. Komaratat, M. Kates, T. Novitsky, and D. J. Kushner. 1975. Isolation and characterization of *Actinopolyspora halophila*, gen. et sp. nov., an extremely halophilic actinomycete. Can. J. Microbiol. 21:1500-1511.
56. Goodfellow, M., L. J. Stanton, K. E. Simpson, and D. E. Minnikin. 1990. Numerical and chemical classification of *Actinoplanes* and some related actinomycetes. J. Gen. Microbiol. 136:19-36.
57. Greenwood, J. R., and M. J. Pickett. 1980. Transfer of *Ilaemophilus vaginalis* Gardner and Dukes to a new genus, *Gardnerella*: *G. vaginalis* (Gardner and Dukes) comb. nov. Int. J. Syst. Bacteriol. 30:170-178.
58. Groth, I., P. Schumann, N. Weiss, K. Martin, and F. A. Rainey. 1996. *Agrococcus jenensis* gen. nov., sp. nov., a new genus of actinomycetes with diaminobutyric acid in the cell wall. Int. J. Syst. Bacteriol. 46:234-239.
59. Gupta, R. S., and G. B. Golding. 1996. The origin of the eukaryotic cell. Trends Biochem. Sci. 21:166-171.
60. Gupta, R. S., and B. Singh. 1992. Cloning of the HSP70 gene from *Halobacterium marismortui*: relatedness of archaeobacterial HSP70 to its eubacterial homologs and a model for the evolution of the HSP70 gene. J. Bacteriol. 174:4594-4605.
61. Haas, F., and H. König. 1988. *Coriobacterium glomerans* gen. nov., spec. nov. from the intestinal tract of the red soldier bug. Int. J. Syst. Bacteriol. 38:382-384.
62. Hahn, D., M. P. Lechevalier, A. Fischer, and E. Stackebrandt. 1989. Evidence for a close phylogenetic relationship between members of the genera *Frankia*, *Geodermatophilus* and *Blastococcus* and emendation of the family *Frankiaceae*. Syst. Appl. Microbiol. 11:236-242.
63. Harz, C. O. 1877. *Actinomyces bovis*, ein neuer Schimmel in den Geweben

- des Rindes. Jahresber. K. Zbl. Thierärz. Sch. München 1877/1888. 5:125-140.
64. Hasegawa, T. 1988. *Actinokineospora*: a new genus of the *Actinomycetales*. *Actinomycetes* 2:31-45.
  65. Hasegawa, T., M. P. Lechevalier, and H. A. Lechevalier. 1978. A new genus of the *Actinomycetales*: *Actinosynnema* gen. nov. *Int. J. Syst. Bacteriol.* 28:304-310.
  66. Henssen, A. 1957. Beiträge zur Morphologie und Systematik der thermophilen Actinomyceten. *Arch. Mikrobiol.* 26:373-416.
  67. Huber, R., C. R. Woese, T. A. Langworthy, H. Fricke, and K. O. Stetter. 1989. *Thermosipho africanus* gen. nov. represents a new genus of thermophilic eubacteria within the "Thermotogales." *Syst. Appl. Microbiol.* 12:32-37.
  68. Iinuma, S., A. Yokota, T. Hasegawa, and T. Kanamaru. 1994. *Actinocorallica* gen. nov., a new genus of the order *Actinomycetales*. *Int. J. Syst. Bacteriol.* 44:230-234.
  69. Jones, D., and M. D. Collins. 1988. Taxonomic studies on some human cutaneous coryneform bacteria: description of *Dermabacter hominus* gen. nov., sp. nov. *FEMS Microbiol. Lett.* 51:51-56.
  70. Jukes, T. H., and C. R. Cantor. 1969. Evolution of protein molecules, p. 21-132. In H. N. Munro (ed.), *Mammalian protein metabolism*. Academic Press, New York, N.Y.
  71. Kalakoutskii, L. V., I. P. Kirillova, and N. A. Krasil'nikov. 1967. A new genus of the *Actinomycetales*—*Intrasporangium* gen. nov. *J. Gen. Microbiol.* 48:79-85.
  72. Kane, W. D. 1966. A new genus of *Actinoplanaceae*, *Pilimelia*, with a description of two species, *Pilimelia terevasa* and *Pilimelia anulata*. *J. Elisha Mitchell Sci. Soc.* 82:220-230.
  73. Kim, D., J. Chun, N. Sahin, Y. C. Hah, and M. Goodfellow. 1996. Analysis of thermophilic clades within the genus *Streptomyces* by 16S ribosomal DNA sequence comparisons. *Int. J. Syst. Bacteriol.* 46:581-587.
  74. Kim, S.-B., J.-H. Yoon, H. Kim, S. T. Lee, Y.-H. Park, and M. Goodfellow. 1995. A phylogenetic analysis of the genus *Saccharomonospora* conducted with 16S rRNA gene sequences. *Int. J. Syst. Bacteriol.* 4:351-356.
  75. Koch, C., R. M. Kroppenstedt, and E. Stackebrandt. 1996. Intrageneric relationships of the actinomycete genus *Micromonospora*. *Int. J. Syst. Bacteriol.* 46:383-387.
  76. Koch, C., F. A. Rainey, and E. Stackebrandt. 1994. 16S rDNA studies on members of *Arthrobacter* and *Micrococcus*, an aid for their future taxonomic restructuring. *FEMS Microbiol. Lett.* 123:167-172.
  77. Koch, C., P. Schumann, and E. Stackebrandt. 1995. Reclassification of *Micrococcus agilis* Ali-Cohen 1889 to the genus *Arthrobacter* as *Arthrobacter agilis* comb. nov. and emendation of the genus *Arthrobacter*. *Int. J. Syst. Bacteriol.* 45:837-839.
  78. Koch, C., R. M. Kroppenstedt, F. A. Rainey, and E. Stackebrandt. 1996. 16S ribosomal DNA analysis of the genera *Micromonospora*, *Actinoplanes*, *Catenulospira*, *Catenuloplanes*, *Couchioplanes*, *Dactylosporangium*, and *Pilimelia* and emendation of the family *Micromonosporaceae*. *Int. J. Syst. Bacteriol.* 46:765-768.
  79. Korn-Wendisch, F., F. A. Rainey, R. M. Kroppenstedt, A. Kempf, A. Majzma, H. J. Kutzner, and E. Stackebrandt. 1995. *Thermocristum* gen. nov., a new genus of the order *Actinomycetales*, and description of *Thermocristum municipale* sp. nov. and *Thermocristum agreste* sp. nov. *Int. J. Syst. Bacteriol.* 45:67-77.
  80. Krasilnikov, N. A. 1938. Ray fungi and related organisms, *Actinomycetales*. *Akademiya Nauk SSSR, Moscow*.
  81. Krasilnikov, N. A. 1949. Guide to the bacteria and actinomycetes. *Akademiya Nauk SSSR Moscow*.
  82. Krasilnikov, N. A., L. V. Kalakoutskii, and N. F. Kirillova. 1961. A new genus of ray fungi—*Promicromonospora* gen. nov. *Izv. Akad. Nauk SSSR Ser. Biol.* 1:107-112.
  83. Kudo, T., T. Itoh, S. Miyadoh, T. Shomura, and A. Seino. 1993. *Herbidospira* gen. nov., a new genus of the family *Streptosporangiaceae* Goodfellow et al. 1990. *Int. J. Syst. Bacteriol.* 43:319-328.
  84. Labeda, D. P., R. T. Testa, M. P. Lechevalier, and H. A. Lechevalier. 1984. *Saccharothrix*, a new genus of the *Actinomycetales* related to *Nocardioopsis*. *Int. J. Syst. Bacteriol.* 34:426-431.
  85. Labeda, D. P., R. T. Testa, M. P. Lechevalier, and H. A. Lechevalier. 1985. *Glycomyces*, a new genus of the *Actinomycetales*. *Int. J. Syst. Bacteriol.* 35:417-421.
  86. Lacey, J., and M. Goodfellow. 1975. A novel actinomycete from sugar cane bagasse: *Saccharopolyspora hirsuta* gen. et sp. nov. *J. Gen. Microbiol.* 88:78-85.
  87. Lassnig, C., M. Dorsch, E. Schaper, H. Stöffler, J. Wolters, and E. Stackebrandt. 1989. Phylogenetic evidence for a relationship of *Mobiluncus* to the genus *Actinomyces*. *FEMS Microbiol. Lett.* 65:17-22.
  88. Leblond-Bourget, N., H. Philippe, I. Mangin, and B. Decaris. 1996. 16S rRNA and 16S to 23S internal transcribed spacer sequence analyses reveal inter- and intraspecific *Bifidobacterium* phylogeny. *Int. J. Syst. Bacteriol.* 46:102-111.
  89. Lechevalier, H. A., and M. P. Lechevalier. 1970. A critical evaluation of the genera of aerobic actinomycetes, p. 393-405. In H. Prauser (ed.), *The Actinomycetales*. VEB Gustav Fischer Verlag, Jena, German Democratic Republic.
  90. Lechevalier, M. P. 1972. Description of a new species, *Oerskovia xanthinolytica*, and emendation of *Oerskovia* Prauser et al. *Int. J. Syst. Bacteriol.* 22:260-264.
  91. Lechevalier, M. P., H. A. Lechevalier, and P. E. Holbert. 1968. *Sporichthya*, un nouveau genre de *Streptomycetaceae*. *Ann. Inst. Pasteur* 114:277-286.
  92. Lechevalier, M. P., H. Prauser, D. P. Labeda, and J.-S. Ruan. 1986. Two new genera of nocardioform actinomycetes: *Amycolata* gen. nov. and *Amycolatopsis* gen. nov. *Int. J. Syst. Bacteriol.* 36:29-37.
  93. Lechevalier, M. P., M. Solotorovsky, and C. I. McDurmont. 1961. A new genus of *Actinomycetales*: *Micropolyspora* gen. nov. *J. Gen. Microbiol.* 26:11-18.
  94. Lehmann, K. B., and R. Neumann. 1896. *Atlas und Grundriss der Bakteriologie und Lehrbuch der speziellen bakteriologischen Diagnostik*, 1st ed. J. F. Lehmann, Munich, Germany.
  95. Lehmann, K. B., and R. Neumann. 1907. *Lehmann's Medizin Handatanten. X. Atlas und Grundriss der Bakteriologie und Lehrbuch der speziellen bakteriologischen Diagnostik*, 4th ed. J. F. Lehmann, Munich, Germany.
  96. Ludwig, W., and K.-H. Schleifer. 1994. Bacterial phylogeny based on 16S and 23S rRNA sequence analysis. *FEMS Microbiol. Rev.* 15:155-173.
  97. Ludwig, W., J. Neumaier, N. Klugbauer, E. Brockmann, C. Roller, S. Jilg, K. Reetz, I. Schachtner, A. Ludvigsen, M. Bachleitner, U. Fischer, and K.-H. Schleifer. 1993. Phylogenetic relationships of *Bacteria* based on comparative sequence analysis of elongation factor Tu and ATP-synthase b-subunit genes. *Antonie Leeuwenhoek* 64:285-305.
  98. Luedemann, G. 1968. *Geodermatophilus*, a new genus of the *Dermatophilaceae* (*Actinomycetales*). *J. Bacteriol.* 96:1848-1858.
  99. Maidak, B. L., N. Larsen, M. J. McCaughey, R. Overbeek, G. J. Olsen, K. Fogel, J. Blandy, and C. R. Woese. 1994. The Ribosomal Database Project. *Nucleic Acids Res.* 22:3485-3487.
  100. Meyer, J. 1976. *Nocardioopsis*, a new genus of the order *Actinomycetales*. *Int. J. Syst. Bacteriol.* 26:487-493.
  101. Miller, E. S., C. R. Woese, and S. Brenner. 1991. Description of the erythromycin-producing bacterium *Arthrobacter* sp. strain NRRL B-3381 as *Acromicrobium erythreum* gen. nov., sp. nov. *Int. J. Syst. Bacteriol.* 41:363-368.
  102. Mohagheghi, A., K. Grohmann, M. Himmel, L. Leighton, and D. M. Updegraff. 1986. Isolation and characterization of *Acidotherrmus cellulolyticus* gen. nov., sp. nov., a new genus of thermophilic, acidophilic, cellulolytic bacteria. *Int. J. Syst. Bacteriol.* 36:435-443.
  103. Nakamura, K., A. Hiraishi, Y. Yoshimi, M. Kawaharasaki, K. Masuda, and Y. Kamagata. 1995. *Microlunatus phosphovorus* gen. nov., sp. nov., a new gram-positive polyphosphate-accumulating bacterium isolated from activated sludge. *Int. J. Syst. Bacteriol.* 45:17-22.
  104. Nesterenko, O., E. I. Krassilnikov, and T. M. Nogina. 1985. *Nocardiodaceae* fam. nov., a new family of the order *Actinomycetales* Buchanan 1917. *Mikrobiol. Zh.* 47:3-12.
  105. Nonomura, H., and Y. Ohara. 1957. Distribution of actinomycetes in soil. II. *Microbispora*, a new genus of *Streptomycetaceae*. *J. Ferment. Technol.* 35:307-311.
  106. Nonomura, H., and Y. Ohara. 1971. Distribution of actinomycetes in soil. X. New genus and species of monosporic actinomycetes. *J. Ferment. Technol.* 49:895-903.
  107. Normand, P., S. Orso, B. Cournoyer, P. Jeannin, C. Chapelon, J. Dawson, L. Evushenko, and A. K. Misra. 1996. Molecular phylogeny of the genus *Frankia* and related genera and emendation of the family *Frankiaceae*. *Int. J. Syst. Bacteriol.* 46:1-9.
  108. Orla-Jensen, S. 1909. Die Hauptlinien des natürlichen Bakteriensystems. *Zentralbl. Bakteriol. Parasitenkd. Infektionskr. Hyg. Abt. 2* 22:305-346.
  109. Orla-Jensen, S. 1919. The lactic acid bacteria. Høst and Son, Copenhagen, Denmark.
  110. Orla-Jensen, S. 1924. La classification des bactéries lactiques. *Lait* 4:468-474.
  111. Ørskov, J. 1923. Investigations into the morphology of the Ray Fungi. Levin and Munksgaard, Copenhagen, Denmark.
  112. Pagani, H., and F. Parenti. 1978. *Kineosporia*, a new genus of the order *Actinomycetales*. *Int. J. Syst. Bacteriol.* 28:401-406.
  113. Park, Y. H., K. Suzuki, D. G. Yim, K. C. Lee, J. Yoon, S. Kim, Y. H. Kho, M. Goodfellow, and K. Komagata. 1993. Suprageneric classification of peptidoglycan group B actinomycetes by sequencing of 5S ribosomal RNA. *Antonie Leeuwenhoek* 64:307-313.
  114. Pascual, C., P. A. Lawson, J. A. E. Farrow, M. N. Gimenez, and M. D. Collins. 1995. Phylogenetic analysis of the genus *Corynebacterium* based on 16S rRNA gene sequences. *Int. J. Syst. Bacteriol.* 45:724-728.
  115. Pitulle, C., M. Dorsch, J. Kazda, J. Wolters, and E. Stackebrandt. 1992. Phylogeny of rapidly growing members of the genus *Mycobacterium*. *Int. J. Syst. Bacteriol.* 42:337-343.
  116. Prauser, H. 1976. *Nocardioidea*, a new genus of the order *Actinomycetales*. *Int. J. Syst. Bacteriol.* 26:58-65.
  117. Prauser, H., M. P. Lechevalier, and H. A. Lechevalier. 1970. Description of

- Oerskovia* gen. n. to harbor Ørskov's motile *Nocardia*. Appl. Microbiol. 19:534.
118. Prévot, A. R. 1940. Manuel de classification et de détermination des bactéries anaérobies, 1st ed. Masson, Paris, France.
  119. Prévot, A. R. 1953. Classification, embranchement, sous-embranchements, classes, ordres, familles, tribus et genres d'après, p. 692. In P. Hauduroy, G. Ehringer, G. Guillot, J. Magrou, A. R. Prévot, A. Rosset, and A. Urbain (ed.), Dictionnaire des bactéries pathogènes, 2nd ed. Masson, Paris, France.
  120. Prévot, A. R. 1961. Traité de systématique bactérienne. Dunod, Paris, France.
  121. Pribham, E. 1929. A contribution to the classification of microorganisms. J. Bacteriol. 18:361–394.
  122. Rainey, F., N. Weiss, H. Prauser, and E. Stackebrandt. 1994. Further evidence for the phylogenetic coherence of actinomycetes with group B-peptidoglycan and evidence for the phylogenetic intermixing of the genera *Microbacterium* and *Aureobacterium* as determined by 16S rRNA analysis. FEMS Microbiol. Lett. 118:135–140.
  123. Rainey, F. A., and E. Stackebrandt. 1993. Phylogenetic evidence for the classification of *Acidothermus cellulolyticus* into the subphylum of actinomycetes. FEMS Microbiol. Lett. 108:27–30.
  124. Rainey, F. A., J. Burghardt, R. M. Kroppenstedt, S. Klatte, and E. Stackebrandt. 1995. Phylogenetic analysis of the genera *Rhodococcus* and *Nocardia* and evidence for the evolutionary origin of the genus *Nocardia* from within the radiation of *Rhodococcus* species. Microbiology (Reading) 141: 523–528.
  125. Rainey, F. A., S. Klatte, R. M. Kroppenstedt, and E. Stackebrandt. 1995. *Dietzia*, a new genus including *Dietzia maris* comb. nov., formerly *Rhodococcus maris*. Int. J. Syst. Bacteriol. 45:32–36.
  126. Rainey, F. A., P. Schumann, H. Prauser, R. Toalster, and E. Stackebrandt. 1993. *Sporichthya polymorpha* represents a novel line of descent within the order *Actinomycetales*. FEMS Microbiol. Lett. 109:263–269.
  127. Rainey, F. A., N. Ward-Rainey, R. M. Kroppenstedt, and E. Stackebrandt. 1996. The genus *Nocardiopsis* represents a phylogenetically coherent taxon and a distinct actinomycete lineage: Proposal of *Nocardiopsaceae* fam. nov. Int. J. Syst. Bacteriol. 46:1088–1092.
  128. Rainey, F. A., N. Weiss, and E. Stackebrandt. 1994. *Coriobacterium* and *Atopobium* are phylogenetic neighbors within the actinomycetes line of descent. Syst. Appl. Microbiol. 17:202–205.
  129. Rainey, F. A., N. Weiss, and E. Stackebrandt. 1995. Phylogenetic analysis of the genera *Cellulomonas*, *Promicromonospora*, and *Jonesia* and proposal to exclude the genus *Jonesia* from the family *Cellulomonadaceae*. Int. J. Syst. Bacteriol. 45:649–652.
  130. Rocourt, J., U. Wehmeyer, and E. Stackebrandt. 1987. Transfer of *Listeria denitrificans* into a new genus *Jonesia* gen. nov. as *Jonesia denitrificans* comb. nov. Int. J. Syst. Bacteriol. 37:266–270.
  131. Rogall, T., J. Wolters, T. Flohr, and E. Böttger. 1990. Towards a phylogeny and the definition of species at the molecular level within the genus *Mycobacterium*. Int. J. Syst. Bacteriol. 40:323–330.
  132. Roller, C., W. Ludwig, and K. H. Schleifer. 1992. Gram-positive bacteria with a high G+C content are characterized by a common insertion within their 23S rRNA genes. J. Gen. Microbiol. 138:1167–1175.
  133. Roller, C., M. Wagner, R. Amann, W. Ludwig, and K.-H. Schleifer. 1994. In situ probing of Gram-positive bacteria with high DNA G+C content using 23S rRNA-targeted oligonucleotides. J. Gen. Microbiol. 140:2859–2865.
  134. Ruimy, R., P. Boiron, V. Boivin, and R. Christen. 1994. A phylogeny of the genus *Nocardia* deduced from the analysis of small-subunit ribosomal DNA sequences, including transfer of *Nocardia amarae* to the genus *Gordonia* as *Gordonia amarae* comb. nov. FEMS Microbiol. Lett. 123:261–268.
  135. Ruimy, R., P. Riegel, P. Boiron, H. Monteil, and R. Christen. 1995. Phylogeny of the genus *Corynebacterium* deduced from analysis of small-subunit ribosomal DNA sequences. Int. J. Syst. Bacteriol. 45:740–746.
  136. Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4:406–425.
  137. Sanders, J. E., and J. L. Fryer. 1980. *Renibacterium salmoninarum* gen. nov., sp. nov., the causative agent of bacterial kidney disease in salmonid fishes. Int. J. Syst. Bacteriol. 30:496–502.
  138. Schlesner, H., and E. Stackebrandt. 1986. Assignment of the genera *Planctomyces* and *Pirella* to a new family *Planctomycetaceae* fam. nov. and description of the order *Planctomycetales* ord. nov. Syst. Appl. Microbiol. 8:174–176.
  139. Schubert, K., W. Ludwig, N. Springer, R. M. Kroppenstedt, J.-P. Accolas, and F. Fiedler. 1996. Two coryneform bacteria isolated from the surface of French Gruyère and Beaufort cheeses are new species of the genus *Brachybacterium*: *Brachybacterium alimentarium* sp. nov. and *Brachybacterium tyrofermentans* sp. nov. Int. J. Syst. Bacteriol. 46:81–87.
  140. Shearer, M. C., P. M. Colman, R. M. Ferrin, L. J. Nisbet, and C. H. Nash III. 1986. A new genus of the *Actinomycetales*: *Kibdelosporangium aridum* gen. nov., sp. nov. Int. J. Syst. Bacteriol. 36:47–54.
  141. Spiegel, C. A., and M. Roberts. 1984. *Mobiluncus* gen. nov., *Mobiluncus curtisii* subsp. *curtisii* sp. nov., *Mobiluncus curtisii* subsp. *holmesii* subsp. nov., and *Mobiluncus mulieris* sp. nov., curved rods from the human vagina. Int. J. Syst. Bacteriol. 34:177–184.
  142. Stackebrandt, E., and O. Charfreitag. 1990. Partial 16S rRNA primary structure of five *Actinomyces* species: phylogenetic implications and development of an *Actinomyces israelii*-specific oligonucleotide probe. J. Gen. Microbiol. 136:37–43.
  143. Stackebrandt, E., and B. M. Goebel. 1994. A place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. Int. J. Syst. Bacteriol. 44:846–849.
  144. Stackebrandt, E., and O. Kandler. 1979. Taxonomy of the genus *Cellulomonas*, based on phylogenetic characters and deoxyribonucleic acid-deoxyribonucleic acid homology, and proposal of seven neotype strains. Int. J. Syst. Bacteriol. 29:273–282.
  145. Stackebrandt, E., and R. Kroppenstedt. 1987. Union of the genera *Actinoplanes* Couch, *Ampullariella* Couch, and *Amorphosporangium* in a redefined genus *Actinoplanes*. Syst. Appl. Microbiol. 9:110–114.
  146. Stackebrandt, E., and W. Ludwig. 1994. The importance of using outgroup reference organisms in phylogenetic studies, the *Atopobium* case. Syst. Appl. Microbiol. 17:39–43.
  147. Stackebrandt, E., and H. Prauser. 1991. Assignment of the genera *Cellulomonas*, *Oerskovia*, *Promicromonospora* and *Jonesia* into *Cellulomonadaceae* family nov. Syst. Appl. Microbiol. 14:261–265.
  148. Stackebrandt, E., B. J. Lewis, and C. R. Woese. 1980. The phylogenetic structure of the coryneform group of bacteria. Zbl. Bakt. Abt. 1 Orig. C 2:137–149.
  149. Stackebrandt, E., W. Liesack, and B. M. Goebel. 1993. Bacterial diversity in a soil sample from a subtropical Australian environment as determined by 16S rDNA analysis. FASEB J. 7:232–236.
  150. Stackebrandt, E., C. Koch, O. Gvozdiak, and P. Schumann. 1995. Taxonomic dissection of the genus *Micromonospora*: *Kocuria* gen. nov., *Nesterenkonia* gen. nov., *Kytococcus* gen. nov., *Dermacoccus* gen. nov., and *Micrococcus* Cohn 1872 gen. emend. Int. J. Syst. Bacteriol. 45:682–692.
  151. Stackebrandt, E., R. Kroppenstedt, and V. Fowler. 1983. A phylogenetic analysis of the family *Dermatophilaceae*. J. Gen. Microbiol. 129:1831–1838.
  152. Stackebrandt, E., R. M. Kroppenstedt, K.-D. Jahnke, C. Kemmerling, and H. Gürtler. 1994. Transfer of *Streptosporangium viridogriseum* (Okuda et al. 1966), *Streptosporangium viridogriseum* subsp. *kofuense* (Nonomura and Ohara 1969), and *Streptosporangium album* (Furumai et al. 1968) to *Kutzneria* gen. nov. as *Kutzneria viridogrisea* comb. nov., *Kutzneria kofuensis* comb. nov., and *Kutzneria alba* comb. nov., respectively, and emendation of the genus *Streptosporangium*. Int. J. Syst. Bacteriol. 44:265–269.
  153. Stackebrandt, E., H. Seiler, and K.-H. Schleifer. 1982. Union of the genera *Cellulomonas* Bergey et al. and *Oerskovia* Prauser et al. in a redefined genus *Cellulomonas*. Zbl. Bakt. Hyg. Abt. 1 Orig. C 3:401–409.
  154. Stackebrandt, E., J. Smida, and M. D. Collins. 1988. Evidence of phylogenetic heterogeneity within the genus *Rhodococcus*, revival of the genus *Gordonia* Tsukamura. J. Appl. Microbiol. 34:341–348.
  155. Stackebrandt, E., D. Witt, C. Kemmerling, R. Kroppenstedt, and W. Liesack. 1991. Designation of streptomycete 16S and 23S rRNA-based target regions for oligonucleotide probes. Appl. Environ. Microbiol. 57:1468–1477.
  156. Strunk, O., and W. Ludwig. 1995. ARB—a software environment for sequence data. Department of Microbiology, Technical University of Munich, Munich, Germany. (E-mail: arb@mikro.biologie.tu-muenchen.de.)
  157. Suzuki, K., M. D. Collins, E. Iijima, and K. Komagata. 1988. Chemotaxonomic characterization of a radiotolerant bacterium, *Arthrobacter radiotolerans*: description of *Rubrobacter radiotolerans* gen. nov., comb. nov. FEMS Microbiol. Lett. 52:33–40.
  158. Suzuki, K. I., J. Sasaki, M. Uramoto, T. Nakase, and K. Komagata. 1996. *Agromyces mediolantis* sp. nov., nom. rev., comb. nov., a species for “*Corynebacterium mediolantis*” Mamoli 1939 and for some aniline-assimilating bacteria which contain 2,4-diaminobutyric acid in the cell wall peptidoglycan. Int. J. Syst. Bacteriol. 46:88–93.
  159. Takeuchi, M., and A. Yokota. 1994. Phylogenetic analysis of the genus *Microbacterium* based on 16S rRNA gene sequences. FEMS Microbiol. Lett. 24:11–16.
  160. Takeuchi, T., H. Sawada, F. Tanaka, and I. Matsuda. 1996. Phylogenetic analysis of *Streptomyces* spp. causing potato scab based on 16S rRNA sequences. Int. J. Syst. Bacteriol. 46:476–479.
  161. Tamura, T., and A. Yokota. 1994. Transfer of *Nocardioides fastidiosum* Collins and Stackebrandt 1989 to the genus *Aeromicrobium fastidiosum* comb. nov. Int. J. Syst. Bacteriol. 44:608–611.
  162. Tamura, T., Y. Nakagaito, T. Nishii, T. Hasegawa, E. Stackebrandt, and A. Yokota. 1994. A new genus of the order *Actinomycetales*, *Couchioplanes* gen. nov., with description of *Couchioplanes caeruleus* Horan and Brodsky 1986 comb. nov. and *Couchioplanes caeruleus* subsp. *azureus* subsp. nov. Int. J. Syst. Bacteriol. 44:193–203.
  163. Tamura, T., M. Takeuchi, and A. Yokota. 1994. *Luteococcus japonicus* gen. nov., sp. nov., a new gram-positive coccus with 11-diaminopimelic acid in the cell wall. Int. J. Syst. Bacteriol. 44:348–356.
  164. Thiemann, J. E., and G. Beretta. 1968. A new genus of the *Actinoplanaceae*. *Planohispora*, gen. nov. Arch. Mikrobiol. 62:157–166.

165. Thiemann, J. E., H. Pagani, and G. Beretta. 1967. A new genus of the *Actinoplanaceae*. *Dactylosporangium*, gen. nov. *Arch. Mikrobiol.* 58:42-52.
166. Thiemann, J. E., H. Pagani, and G. Beretta. 1967. A new genus of the *Actinoplanaceae*. *Planomonospora*, gen. nov. *Giornale Microbiol.* 15:27-28.
167. Thiemann, J. E., H. Pagani, and G. Beretta. 1968. A new genus of *Actinomycetales*. *Microtetraspora*, gen. nov. *J. Gen. Microbiol.* 50:295-303.
168. Tomita, K., Y. Nakakita, Y. Hoshino, K. Numata, and H. Kawaguchi. 1987. New genus of the *Actinomycetales*: *Streptoalloteichus hindustanus* gen. nov., nom. rev.; sp. nov., nom. rev. *Int. J. Syst. Bacteriol.* 37:211-213.
169. Trevisan, V. 1889. I generi e le specie delle Batteriaceae. Zanaboni and Gabuzzi, Milan, Italy.
170. Van Saceghem, R. 1915. Dermatoze contagieuse (impétigo contagieux). *Bull. Soc. Pathol. Exot.* 8:354-359.
171. Waksman, S. A., and A. T. Henrici. 1943. The nomenclature and classification of the actinomycetes. *J. Bacteriol.* 46:337-341.
172. Wang, Y., Z. Zhang, and J. Ruan. 1996. Phylogenetic analysis reveals new relationships among members of the genera *Microtetraspora* and *Microbispora*. *Int. J. Syst. Bacteriol.* 46:658-663.
173. Ward-Rainey, N., F. A. Rainey, H. Schlesner, and E. Stackebrandt. 1995. Assignment of hitherto unidentified 16S rDNA species to a main line of descent within the domain *Bacteria*. *Microbiology* 141:3247-3250.
174. Ward-Rainey, N., F. A. Rainey, and E. Stackebrandt. 1996. The phylogenetic structure of the genus *Streptosporangium*. *Syst. Appl. Microbiol.* 19: 50-55.
175. Warwick, S., T. Bowen, H. McVeigh, and T. M. Embley. 1994. A phylogenetic analysis of the family *Pseudonocardiaceae* and the genera *Actinokinetospora* and *Saccharothrix* with 16S rRNA sequences and a proposal to combine the genera *Amycolata* and *Pseudonocardia* in an emended genus *Pseudonocardia*. *Int. J. Syst. Bacteriol.* 44:293-299.
176. Wellington, E. M. H., E. Stackebrandt, D. Sanders, J. Wolstrup, and N. O. G. Jorgensen. 1992. Taxonomic status of *Kilusatospora*, and proposed unification with *Streptomyces* on the basis of phenotypic and 16S rRNA analysis and emendation of *Streptomyces* Waksman and Henrici 1943<sup>AL</sup>. *Int. J. Syst. Bacteriol.* 42:156-160.
177. Wheelis, M. L., O. Kandler, and C. R. Woese. 1992. On the nature of global classification. *Proc. Natl. Acad. Sci. USA* 89:2930-2934.
178. Witt, D., and E. Stackebrandt. 1990. Unification of the genera *Streptovericillium* and *Streptomyces*, and amendment of *Streptomyces* Waksman and Henrici 1943, 339<sup>AL</sup>. *Syst. Appl. Microbiol.* 13:361-371.
179. Woese, C. R. 1987. Bacterial evolution. *Microbiol. Rev.* 51:221-271.
180. Woese, C. R., O. Kandler, and M. L. Wheelis. 1990. Towards a natural system of organisms. Proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci. USA* 87:4576-4579.
181. Yamada, K., and K. Komagata. 1972. Taxonomic studies on coryneform bacteria. IV. Morphological, cultural, biochemical and physiological characteristics. *J. Gen. Appl. Microbiol.* 18:399-416.
182. Yamamoto, N., S.-I. Sato, K. Saito, T. Hasuo, M. Tadenuma, K.-I. Suzuki, J. Tamaoka, and K. Komagata. 1988. *Rarobacter faecitabidus* gen. nov., sp. nov., a yeast-lysing coryneform bacterium. *Int. J. Syst. Bacteriol.* 38:7-11.
183. Yassin, A. F., F. A. Rainey, H. Brzezinka, J. Burghardt, M. Rafai, P. Seifert, K. Feldmann, and K. P. Schaal. 1996. *Tsukamurella pulmonis* sp. nov. *Int. J. Syst. Bacteriol.* 46:429-436.
184. Yassin, A. F., F. A. Rainey, H. Brzezinka, K.-D. Jahnke, H. Weissbrodt, H. Budzikiewicz, E. Stackebrandt, and K. P. Schaal. 1995. *Lenizea* gen. nov., a new genus of the order *Actinomycetales*. *Int. J. Syst. Bacteriol.* 45:357-363.
185. Yokota, A., T. Tamura, T. Hasegawa, and L. H. Huang. 1993. *Catenuloplanes japonicus* gen. nov., sp. nov., nom. rev., a new genus of the order *Actinomycetales*. *Int. J. Syst. Bacteriol.* 43:805-812.
186. Yokota, A., T. Tamura, T. Nishii, and T. Hasegawa. 1993. *Kineococcus aurantiacus* gen. nov., sp. nov., a new aerobic, gram-positive, motile coccus with meso-diaminopimelic acid and arabinogalactan in the cell wall. *Int. J. Syst. Bacteriol.* 43:52-57.
187. Yokota, A., T. Tamura, N. Weiss, and E. Stackebrandt. 1994. Transfer of *Propionibacterium innocuum* Pitcher and Collins 1991 to *Propioniferax* gen. nov. as *Propioniferax innocuum* comb. nov. *Int. J. Syst. Bacteriol.* 44:579-582.
188. Yoshimi, Y., A. Hiraishi, and K. Nakamura. 1996. Isolation and characterization of *Micrusphaera multipartita* gen. nov., sp. nov., a polysaccharide-accumulating gram-positive bacterium from activated sludge. *Int. J. Syst. Bacteriol.* 46:519-525.
189. Zgurskaya, H. I., L. I. Evtushenko, V. N. Akimov, and L. V. Kalakoutskii. 1993. *Rathayibacter* gen. nov., including the species *Rathayibacter rathayi* comb. nov., *Rathayibacter tritici* comb. nov., *Rathayibacter iranicus* comb. nov., and six strains from annual grasses. *Int. J. Syst. Bacteriol.* 43:143-149.
190. Zopf, W. 1891. Über Ausscheidung von Fettfarbstoffen Lipochromen seitens gewisser Spaltpilze. *Ber. Dtsch. Bot. Ges.* 9:22-29.

Corrected Version NEREUS.079VPC

PATENT COOPERATION TREATY

From the INTERNATIONAL SEARCHING AUTHORITY

PCT

To:

KNOBBE, MARTENS, OLSON & BEAR, LLP  
Attn. Hart, Daniel  
2040 Main Street  
14th Floor  
Irvine, CA 92614  
UNITED STATES OF AMERICA

COMMUNICATION IN CASES FOR WHICH  
NO OTHER FORM IS APPLICABLE

Date of mailing  
(day/month/year)

05/04/2005

Applicant's or agent's file reference

NEREUS.079VP

REPLY DUE

See paragraph 1 below

International application No.

PCT/US2004/019543

International filing date  
(day/month/year)

18/06/2004

Applicant

NEREUS PHARMACEUTICALS, INC.

1. ☐ REPLY DUE within \_\_\_\_\_ ~~XXXX~~ days from the above date of mailing

☒ NO REPLY DUE

2. COMMUNICATION:

Referring to your fax of 01.04.2005, please find enclosed a corrected version of the International Search Report, that replaces and cancels the one already in your possession.

A copy of this letter has been sent to the International Bureau.

Name and mailing address of the International Searching Authority



European Patent Office, P.B. 5818 Patentlaan 2  
NL-2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Joëlle Gerber

# PATENT COOPERATION TREATY

From the INTERNATIONAL SEARCHING AUTHORITY

# PCT

NOTIFICATION OF TRANSMITTAL OF  
THE INTERNATIONAL SEARCH REPORT AND  
THE WRITTEN OPINION OF THE INTERNATIONAL  
SEARCHING AUTHORITY, OR THE DECLARATION

To:  
KNOBBE, MARTENS, OLSON & BEAR, LLP  
Attn. Hart, Daniel  
2040 Main Street  
14th Floor  
Irvine, CA 92614  
UNITED STATES OF AMERICA

(PCT Rule 44.1)

Date of mailing  
(day/month/year) - 5 04 2005

Applicant's or agent's file reference  
NEREUS.079VP

FOR FURTHER ACTION See paragraphs 1 and 4 below

International application No.  
PCT/US2004/019543

International filing date  
(day/month/year) 18/06/2004

Applicant  
NEREUS PHARMACEUTICALS, INC.

1. ☒ The applicant is hereby notified that the international search report and the written opinion of the International Searching Authority have been established and are transmitted herewith.

## Filing of amendments and statement under Article 19:

The applicant is entitled, if he so wishes, to amend the claims of the International Application (see Rule 46):

**When?** The time limit for filing such amendments is normally 2 months from the date of transmittal of the International Search Report; however, for more details, see the notes on the accompanying sheet.

**Where?** Directly to the International Bureau of WIPO, 34 chemin des Colombettes  
1211 Geneva 20, Switzerland, Facsimile No.: (41-22) 740.14.35

For more detailed instructions, see the notes on the accompanying sheet.

2. ☐ The applicant is hereby notified that no international search report will be established and that the declaration under Article 17(2)(a) to that effect and the written opinion of the International Searching Authority are transmitted herewith.

3. ☐ With regard to the protest against payment of (an) additional fee(s) under Rule 40.2, the applicant is notified that:

- ☐ the protest together with the decision thereon has been transmitted to the International Bureau together with the applicant's request to forward the texts of both the protest and the decision thereon to the designated Offices.  
☐ no decision has been made yet on the protest; the applicant will be notified as soon as a decision is made.

## 4. Reminders

Shortly after the expiration of 18 months from the priority date, the international application will be published by the International Bureau. If the applicant wishes to avoid or postpone publication, a notice of withdrawal of the international application, or of the priority claim, must reach the International Bureau as provided in Rules 90bis.1 and 90bis.3, respectively, before the completion of the technical preparations for international publication.

The applicant may submit comments on an informal basis on the written opinion of the International Searching Authority to the International Bureau. The International Bureau will send a copy of such comments to all designated Offices unless an international preliminary examination report has been or is to be established. These comments would also be made available to the public but not before the expiration of 30 months from the priority date.

Within 19 months from the priority date, but only in respect of some designated Offices, a demand for international preliminary examination must be filed if the applicant wishes to postpone the entry into the national phase until 30 months from the priority date (in some Offices even later); otherwise, the applicant must, within 20 months from the priority date, perform the prescribed acts for entry into the national phase before those designated Offices.

In respect of other designated Offices, the time limit of 30 months (or later) will apply even if no demand is filed within 19 months.

See the Annex to Form PCT/IB/301 and, for details about the applicable time limits, Office by Office, see the *PCT Applicant's Guide*, Volume II, National Chapters and the WIPO Internet site.

Name and mailing address of the International Searching Authority



European Patent Office, P.B. 5818 Patentlaan 2  
NL-2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax (+31-70) 340-2016

Authorized officer

Joëlle Gerber

## PATENT COOPERATION TREATY

## PCT

## INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference	<b>FOR FURTHER ACTION</b>		see Form PCT/ISA/220 as well as, where applicable, item 5 below.
NEREUS.079VP	International filing date (day/month/year)	(Earliest) Priority Date (day/month/year)	
International application No.	18/06/2004	20/06/2003	
PCT/US2004/019543			
Applicant			
NEREUS PHARMACEUTICALS, INC.			

This International Search Report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This International Search Report consists of a total of 11 sheets.



It is also accompanied by a copy of each prior art document cited in this report.

## 1. Basis of the report

- a. With regard to the language, the international search was carried out on the basis of the international application in the language in which it was filed, unless otherwise indicated under this item.



The international search was carried out on the basis of a translation of the international application furnished to this Authority (Rule 23.1(b)).



b. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, see Box No. I.

2.



Certain claims were found unsearchable (See Box II).

3.



Unity of invention is lacking (see Box III).

## 4. With regard to the title,



the text is approved as submitted by the applicant.



the text has been established by this Authority to read as follows:

USE OF [3.2.0] HETEROCYCLIC COMPOUNDS AND ANALOGS THEREOF FOR THE TREATMENT OF  
CANCER, INFLAMMATION AND INFECTIOUS DISEASES

## 5. With regard to the abstract,



the text is approved as submitted by the applicant.



the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box No. IV. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.

## 6. With regard to the drawings,



a. the figure of the drawings to be published with the abstract is Figure No. 1



as suggested by the applicant.



as selected by this Authority, because the applicant failed to suggest a figure.



as selected by this Authority, because this figure better characterizes the invention.



b. none of the figures is to be published with the abstract.

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2004/019543

Box No. IV Text of the abstract (Continuation of item 5 of the first sheet)

Disclosed are methods of treating cancer, inflammatory conditions, and/or infectious disease in an animal comprising: administering to the animal, a therapeutically effective amount of a [3.2.0] heterocyclic compound, preferably a 6-oxa-2-azabicyclo[3.2.0]heptane-3,7-dione, such as salinosporamide A. The animal is a mammal, preferably a human or a rodent.



## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US2004/019543**Box II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)**

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:  
  
Although claims 1-14, and 17-24 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
2. ☐ Claims Nos.:  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:  
because they are dependant claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☒ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☒ No protest accompanied the payment of additional search fees.

## FURTHER INFORMATION CONTINUED FROM PCT/ISA/210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1 (partially), 2-15 (entirely), 17-20 (entirely)

Use of a compound having the structure of formula I, and pharmaceutically acceptable salts and pro-drug esters thereof in the treatment of cancer.

A method of treating a neoplastic disease in an animal, the method comprising administering to the animal a therapeutically effective amount of a compound of a formula selected from formulae I-V, and pharmaceutically acceptable salts and pro-drug esters thereof.

A pharmaceutical composition comprising a compound of a formula selected from formulae I-V, and pharmaceutically acceptable salts and pro-drug esters thereof, without an anti-microbial agent.

A method of inhibiting the growth of a cancer cell comprising contacting a cancer cell with a compound of a formula selected from formulae I-V, and pharmaceutically acceptable salts and pro-drug esters thereof.

A method of inhibiting proteasome activity or of inhibiting NF-kappaB activation comprising the step contacting a cell with a compound of a formula selected from formulae I-V, and pharmaceutically acceptable salts and pro-drug esters thereof.

---

2. claims: 1 (partially), 21-22 (entirely)

Use of a compound having the structure of formula I, and pharmaceutically acceptable salts and pro-drug esters thereof in the treatment of inflammation.

A method for treating an inflammatory condition comprising administering an effective amount of a compound of a formula selected from formulae I-V to a patient in need thereof.

---

3. claims: 1 (partially), 16 (entirely), 23-24 (entirely)

Use of a compound having the structure of formula I, and pharmaceutically acceptable salts and pro-drug esters thereof in the treatment of infectious disease.

A pharmaceutical composition comprising a compound of a formula selected from formulae I-V, and pharmaceutically acceptable salts and pro-drug esters thereof and an anti-microbial agent.

A method for treating a microbial illness comprising administering an effective amount of a compound of a formula selected from formulae I-V to a patient in need thereof.

---

# INTERNATIONAL SEARCH REPORT

CORRECTED VERSION

International Application No  
PCT/US2004/019543

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 A61K31/407 A61P35/00 A61P29/00 A61P19/02  
A61P11/06 A61P25/00 A61P17/06 A61P9/10 A61P33/00  
A61P33/02 A61P33/06

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ, CHEM ABS Data, EMBASE, MEDLINE, SCISEARCH, CANCERLIT  
NAPRALERT, BIOSIS

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	FELING ROBERT H ET AL: "Salinosporamide A: a highly cytotoxic proteasome inhibitor from a novel microbial source, a marine bacterium of the new genus salinospora." ANGEWANDTE CHEMIE (INTERNATIONAL ED. IN ENGLISH), (2003 JAN 20) 42 (3) 355-7. JOURNAL CODE: 0370543. ISSN: 0570-0833., 20 January 2003 (2003-01-20), XP002304842 page 355, column 1, paragraph 3 - column 2, paragraph 1	1-5, 9-12,14, 15,17-20
Y	page 356, column 2, paragraph 3 ----- -/-	1-3,6,8, 16,20-24

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

### \* Special categories of cited documents :

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

\*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

\*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

\*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

\*G\* document member of the same patent family

Date of the actual completion of the international search

18 February 2005

Date of mailing of the international search report

- 5. 04 2005

Name and mailing address of the ISA  
European Patent Office, P.B. 5818 Patentkan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax (+31-70) 340-2016

Authorized officer

Cielen, E

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/US2004/019543

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	FENICAL W. ET AL: "Marine microorganisms as a developing resource for drug discovery." PHARMACEUTICAL NEWS, 9/6 (489-494). REFS: 14 ISSN: 1071-894X CODEN: PHNEEP, 2002, XP008038137	1-3, 9-12,14, 15,17, 19,20
Y	page 493, column 1, paragraph 1  figure 4	1-3,6,8, 16,20-24
X	----- WO 02/47610 A (UNIV CALIFORNIA) 20 June 2002 (2002-06-20)	1-4, 9-12,14, 15,17, 19,20
A	page 9, paragraph 2 - page 10, paragraph 1 examples 5-8,12 claims 1,3-6 figure 1	21,23
X	----- [Online] XP002304843 Retrieved from the Internet: URL:http://www.newton.rcs.it/PrimoPiano/News/2003/02_Febbraio/03/Antobiotico.shtml> [retrieved on 2003-02-02]	1-4, 9-12,14, 15,17, 19,20
A	the whole document	23
Y	----- GOLDBERG, ALFRED L. ET AL: "Not just research tools- proteasome inhibitors offer therapeutic promise" NATURE MEDICINE (NEW YORK, NY, UNITED STATES) , 8(4), 338-340 CODEN: NAMEFI; ISSN: 1078-8956, 2002, XP008038140 page 338, column 2, paragraph 1 - column 3, paragraph 1 page 340, column 2, paragraph 2 - paragraph 3	1-3,6,8, 20-22
A	----- NICOLAUS B J R: "Symbiotic Approach to Drug Design" DECISION MAKING IN DRUG RESEARCH, XX, XX, 1983, pages 173-186, XP002197412	
P,X	----- REDDY, LELETI RAJENDER ET AL: "A Simple Stereocontrolled Synthesis of Salinosporamide A" JOURNAL OF THE AMERICAN CHEMICAL SOCIETY , 126(20), 6230-6231 CODEN: JACSAT; ISSN: 0002-7863, 22 April 2004 (2004-04-22), XP008038141	1-3, 9-12,14, 15,17, 19,20
P,Y	page 6230, column 1, paragraph 1  page 6231, column 2, paragraph 2  ----- -/--	1-3,6,8, 16,20-24

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/US2004/019543

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
E	<p>WO 2004/071382 A (BAYER HEALTHCARE AG ; MAYER-BARTSCHMID ANKE (DE); MUELLER HARTWIG (DE)) 26 August 2004 (2004-08-26) page 1, paragraph 1 page 3, paragraph 2 - page 4, paragraph 2 page 6, paragraph 3 - page 12, paragraph 1 page 19, paragraph 6 - paragraph 3 examples 1-7 page 56, paragraph 4 - page 57, paragraph 1 page 58, paragraph 5 - page 60, paragraph 1 claims 1-6,8,9,11-14</p> <p>-----</p>	1-5, 9-12, 14-23
E	<p>US 2004/138196 A1 (FELING ROBERT H R ET AL) 15 July 2004 (2004-07-15) page 1, paragraphs 4,10-16,18 page 3, paragraph 49-51 page 4, paragraphs 56,58 page 4, paragraph 61 - page 5, paragraph 64 claims 1-8,10-21</p> <p>-----</p>	1-12, 14-20,23
Y	<p>WO 96/32105 A (PRESIDENT AND FELLOWS OF HARVARD COLLEGE; SCHREIBER, STUART, L; STANDA) 17 October 1996 (1996-10-17) page 3, line 33 - page 6, line 15 page 7, line 21 - page 8, line 26 page 13, line 27 - page 14, line 12 page 81, lines 1-14 page 82, line 19 - page 83, line 2 page 84, lines 8-22 page 85, lines 3-32 page 87, lines 8-29 page 88, line 12 - page 89, line 14 page 90, line 16 - page 91, line 14 page 91, line 31 - page 92, line 4 claims 6-8,10,23,32,38,40,45-48,53-55,58</p> <p>-----</p>	1-5, 9-12,14, 16-23
Y	<p>WO 00/23614 A (LEUKOSITE, INC; MILLENNIUM PHARMACEUTICALS, INC) 27 April 2000 (2000-04-27) page 3, lines 3,4 page 7, lines 6-24 page 10, lines 14-23 page 11, lines 9-14</p> <p>-----</p>	1,19-22
Y	<p>US 2001/002391 A1 (BRAND STEPHEN ET AL) 31 May 2001 (2001-05-31) page 1, paragraph 10 page 2, paragraphs 14,25,26 page 3, paragraph 29 page 4, paragraph 51 page 5, paragraphs 74,76 claims 1-3,7,11,14,16,20</p> <p>-----</p>	1,19-22

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/US2004/019543

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2001/051654 A1 (ELLIOTT PETER J ET AL) 13 December 2001 (2001-12-13) page 1, paragraph 3 page 2, paragraph 14-17 page 5, paragraph 76 figures scheme,1 claims 1,3-5,7,9-11 -----	1,19-22
Y	ELLIOTT, PETER J. ET AL: "The proteasome : A new target for novel drug therapies" AMERICAN JOURNAL OF CLINICAL PATHOLOGY , 116(5), 637-646 CODEN: AJCPAI; ISSN: 0002-9173, 2001, XP008007324 abstract table 1 page 638, column 2, paragraph 1 page 639, column 2, paragraph 3 page 639, column 2, paragraph 5 - page 640, column 1, paragraph 2 page 641, column 1, paragraph 1 - page 642, column 2, paragraph 4 -----	1-5, 9-12,14, 17-23
Y	CRANE, SHELDON N. ET AL: "A Novel Enantioselective Synthetic Route to Omuralide Analogues with the Potential for Species Selectivity in Proteasome Inhibition" ORGANIC LETTERS , 3(9), 1395-1397 CODEN: ORLEF7; ISSN: 1523-7060, 2001, XP008038134 page 1395, column 1, paragraph 1 - page 1396, column 2, paragraph 1 -----	1,23,24
Y	GANTT, SOREN M. ET AL: "Proteasome inhibitors block development of Plasmodium spp." ANTIMICROBIAL AGENTS AND CHEMOTHERAPY , 42(10), 2731-2738 CODEN: AMACCQ; ISSN: 0066-4804, 1998, XP008038135 abstract page 2731, column 2, paragraph 2 figure 5 page 2736, column 1, paragraph 2 page 2736, column 2, paragraph 3 page 2737, column 1, paragraph 3 ----- -/--	1,23,24

# INTERNATIONAL SEARCH REPORT

CORRECTED VERSION

International Application No

PCT/US2004/019543

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	<p>TANG GUANGQING ET AL: "Proteasome activity is required for anthrax lethal toxin to kill macrophages"</p> <p>INFECTION AND IMMUNITY, vol. 67, no. 6, June 1999 (1999-06), pages 3055-3060, XP008038136</p> <p>ISSN: 0019-9567</p> <p>abstract</p> <p>page 3055, column 2, paragraph 3</p> <p>figure 1</p> <p>page 3056, column 2, paragraphs 2,3</p> <p>page 3058, column 1, last paragraph - column 2, paragraph 1</p> <p>-----</p>	1,23,24
Y	<p>BEERS M., BERKOW R.: "The Merck Manual of Diagnosis and Therapy, seventeenth edition"</p> <p>1999, MERCK RESEARCH LABORATORIES, WHITEHOUSE STATION N.J., XP002318189</p> <p>page 1157, column 2, paragraph 5 - page 1158, column 2, last paragraph</p> <p>page 1241, column 2, paragraph 2 - page 1252, column 1, last paragraph</p> <p>-----</p>	16
E	<p>WO 2005/003137 A (THE REGENTS OF THE UNIVERSITY OF CALIFORNIA; FENICAL, WILLIAM, H; JENS)</p> <p>13 January 2005 (2005-01-13)</p> <p>page 1, paragraph 3</p> <p>page 2, paragraph 7 - page 7, paragraph 14</p> <p>page 7, paragraph 16 - page 8, paragraph 18</p> <p>page 9, paragraph 27 - page 13, paragraph 32</p> <p>page 16, paragraph 49 - page 17, paragraph 51</p> <p>page 19, paragraph 59</p> <p>page 20, paragraph 61 - page 22, paragraph 65</p> <p>example 2</p> <p>claims 1,7-14</p> <p>-----</p>	1-12, 14-20,23

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US2004/019543

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 0247610	A	20-06-2002	AU 4322802 A CA 2429163 A1 EP 1341414 A2 JP 2004535766 T WO 0247610 A2 US 2003157695 A1	24-06-2002 20-06-2002 10-09-2003 02-12-2004 20-06-2002 21-08-2003
WO 2004071382	A	26-08-2004	WO 2004071382 A2	26-08-2004
US 2004138196	A1	15-07-2004	US 2004259856 A1 WO 2005003137 A1	23-12-2004 13-01-2005
WO 9632105	A	17-10-1996	US 6335358 B1 AU 705791 B2 AU 5542396 A CA 2217817 A1 EP 0820283 A1 JP 11503732 T NZ 306775 A WO 9632105 A1 US 5756764 A US 6147223 A US 6458825 B1 US 2003119887 A1 US 6214862 B1 US 6645999 B1 ZA 9602933 A	01-01-2002 03-06-1999 30-10-1996 17-10-1996 28-01-1998 30-03-1999 28-10-1999 17-10-1996 26-05-1998 14-11-2000 01-10-2002 26-06-2003 10-04-2001 11-11-2003 03-02-1997
WO 0023614	A	27-04-2000	AT 278803 T AU 762373 B2 AU 6433399 A BR 9914648 A CA 2347275 A1 CN 1331752 T DE 69920940 D1 EP 1123412 A1 JP 2002527114 T NZ 511197 A WO 0023614 A1 US 6613541 B1	15-10-2004 26-06-2003 08-05-2000 27-11-2001 27-04-2000 16-01-2002 11-11-2004 16-08-2001 27-08-2002 29-08-2003 27-04-2000 02-09-2003
US 2001002391	A1	31-05-2001	US 2003069189 A1 US 2001053760 A1 AT 240113 T AU 741802 B2 AU 6171998 A CA 2281224 A1 CN 1250376 T DE 69814618 D1 DE 69814618 T2 EP 0979096 A1 ES 2196547 T3 HK 1025266 A1 JP 2001511814 T NO 993935 A NZ 337364 A PT 979096 T WO 9835691 A1	10-04-2003 20-12-2001 15-05-2003 13-12-2001 08-09-1998 20-08-1998 12-04-2000 18-06-2003 25-03-2004 16-02-2000 16-12-2003 05-09-2003 14-08-2001 04-10-1999 29-06-2001 30-09-2003 20-08-1998



## INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US2004/019543

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2001051654	A1	13-12-2001	AU 9580098 A 12-04-1999
			CA 2304622 A1 01-04-1999
			EP 1017398 A1 12-07-2000
			JP 2001517631 T 09-10-2001
			WO 9915183 A1 01-04-1999
-----			
WO 2005003137	A	13-01-2005	US 2004138196 A1 15-07-2004
			US 2004259856 A1 23-12-2004
			WO 2005003137 A1 13-01-2005
-----			

**WRITTEN OPINION OF THE  
INTERNATIONAL SEARCHING  
AUTHORITY (SEPARATE SHEET)**

International application No.

PCT/US2004/022242

**Re Item III.**

Claims 1-13 relate to subject-matter considered by this Authority to be covered by the provisions of Rule 67.1(iv) PCT. Consequently, no opinion will be formulated with respect to the industrial applicability of the subject-matter of these claims (Article 34(4)(a)(I) PCT).

**Re Item IV.**

The separate inventions/groups of inventions are:

1. Claims: 1(partially),2-9,10(partially),11-21,22(partially)  
Compound of formula I wherein R40 is a moiety of formula II for inhibiting angiogenesis and treating a tumor.  
Compound of formula V for inhibiting angiogenesis and treating a tumor.
2. Claims: 1,10,22 (all partially)  
Compound of formula I wherein R40 is a 4- or 5-membered, N-containing heterocyclic ring for inhibiting angiogenesis and treating a tumor
3. Claims: 1(partially),10(partially),23  
Compound of formula IV for inhibiting angiogenesis and treating a tumor
4. Claims: 1(partially),10(partially),24  
Compound of formula IX for inhibiting angiogenesis and treating a tumor

They are not so linked as to form a single general inventive concept (Rule 13.1 PCT) for the following reasons:

The problem to be solved by the present invention is to provide a medicament for inhibiting angiogenesis and treating tumors.

The proposed solution is to use thalidomide analogs in particular:

1. a) Compounds of formula (I) wherein R40 is a moiety of formula (II) and compounds of formula (V)  
b) Compounds of formula (I) wherein R40 is a 4- or 5-membered, N-containing heterocyclic ring
2. Compounds of formula (IV)
3. Compounds of formula (IX)

Where claims define chemical alternatives, unity of invention should be considered to be present when the alternatives are of a similar nature [compare "Administrative Instructions under the PCT", Annex B, Unity of Invention, paragraph (f)].

Alternatives chemical compounds are to be regarded as being of a similar nature where:

(i) all alternatives have a common property or activity

**and**

(ii) a common structure is present, i.e. a significant structural element is shared by all of the alternatives, or in case a common structure is absent, all alternatives belong to a recognised class of chemical compounds in the art to which the invention pertains.

The three alternatives 1,2 and 3 claimed do not share a significant structural element (their structure are markedly different) and do not belong to a recognised group or a class of compounds, which may be expected to behave in the same way in the context of the claimed inventions.

Furthermore, the fact that the different alternatives 1a, 1b, 2 and 3 can be considered as thalidomide analogs cannot represent the single general inventive concept linking the different compounds 1a,1b,2 and 3, since thalidomide and certain of its analogs (including compounds of formula (I) and (V)) have been already described as antiangiogenic agents in the state of the art.

The document XP1202664 discloses the antiangiogenic and antitumor activities of the

**WRITTEN OPINION OF THE  
INTERNATIONAL SEARCHING  
AUTHORITY (SEPARATE SHEET)**

International application No.

PCT/US2004/022242

thalidomide analogs of formula (V), CPS13, CPS16 and CPS20.

WO03/014315 discloses the antiangiogenic and antitumor activities of 3-amino-thalidomide and 2,6-diamino-thalidomide (compounds of formula (I)).

WO02/064083 and XP2959459 disclose the antiangiogenic and antitumor activities of 3-amino-thalidomide (compound of formula (I)).

In the present application no further technical feature(s) can be distinguished that can be regarded as a "special technical feature" involved in the technical relationship among the different inventions. Consequently, the present application lacks unity of invention, and the different solutions not belonging to a common inventive concept are identified as the different subjects listed below. Each of the inventions listed is a distinct invention, characterised by its own special technical feature, defining the contribution which each of the claimed inventions, considered as a whole, makes over the prior art, i.e. the specific features of the individual group of compounds.

In reply to the objection to lack of unity, the applicant has paid all the additional search fees.

Consequently, this report has been established in respect of all parts of the international application.

**Re Item V.**

- 1 The following document is referred to in this communication:

D1: NG, SYLVIA S. W. ET AL: "Antiangiogenic Activity of N-substituted and Tetrafluorinated Thalidomide Analogues" CANCER RESEARCH , 63(12), 3189-3194 CODEN: CNREA8; ISSN: 0008-5472, 15 June 2003 (2003-06-15), XP001202664

**WRITTEN OPINION OF THE  
INTERNATIONAL SEARCHING  
AUTHORITY (SEPARATE SHEET)**

International application No.

PCT/US2004/022242

- D2: WO 03/014315 A (ENTREMED INC ; CHILDRENS MEDICAL CENTER (US))  
20 February 2003 (2003-02-20)
- D3: WO 02/064083 A (ROUGAS JOHN ; CONNER BARRY P (US); PRIBLUDA  
VICTOR (US); SHAH JAMSHED) 22 August 2002 (2002-08-22)
- D4: LENTZSCH S ET AL: "S-3-amino-phthalimido-glutarimide inhibits angiogenesis  
and growth of B-cell neoplasias in mice" CANCER RESEARCH, AMERICAN  
ASSOCIATION FOR CANCER RESEARCH, BALTIMORE, MD, US, vol. 62, 15  
April 2002 (2002-04-15), pages 2300-2305, XP002959459 ISSN: 0008-5472
- D5: HESS, S. ET AL: "Synthesis and immunological activity of water-soluble  
thalidomide prodrugs" BIOORGANIC & MEDICINAL CHEMISTRY , 9(5), 1279-  
1291 CODEN: BMECEP; ISSN: 0968-0896, 2001, XP001202654
- D6: DE A U ET AL: "POSSIBLE ANTINEOPLASTIC AGENTS: PART IV -  
SYNTHESIS & ANTINEOPLASTIC POTENCY OF N-SUBSTITUTED ALPHA-  
(4,5-DIMETHOXYPHTHALIMIDO)GLUTARIMIDES & N-SUBSTITUTED BETA-  
(4-BROMOPHENYL)GLUTARIMIDES" INDIAN JOURNAL OF CHEMISTRY,  
SECTION B: ORGANIC, INCL. MEDICINAL, PUBLICATIONS &  
INFORMATIONS DIRECTORATE, NEW DELHI, IN, vol. 16B, no. 6, 1 June  
1978 (1978-06-01), pages 510-512, XP000675183 ISSN: 0019-5103
- D7: DE A U ET AL: "POSSIBLE ANTINEOPLASTIC AGENTS: III SYNTHESIS OF  
6-ALKYL-2- U4'METHOXYPHTHALIMIDO AND 6-ALKYL-3-U3'-4'-  
DIMETHOXYPHENYL GLUTARIMIDES" JOURNAL OF THE INDIAN  
CHEMICAL SOCIETY, THE INDIAN CHEMICAL SOCIETY, CALCUTTA, IN,  
vol. 53, no. 11, 1 November 1976 (1976-11-01), pages 1122-1125,  
XP000675187 ISSN: 0019-4522
- D8: WO 94/20085 A (CHILDRENS HOSP MEDICAL CENTER) 15 September 1994  
(1994-09-15)
- D9: GB-A-1 075 420 (CHEMIE GRUNENTHAL G.M.B.H) 12 July 1967
- D10: US-A-3 560 495 (ERNST FRANKUS ET AL) 2 February 1971 (1971-02-02)
- D11: WO 98/25895 A (ELI LILLY AND COMPANY; ANDERSON, BENJAMIN, A;  
BECKER, GERALD, W; CARTY) 18 June 1998
- D12: GB 962 857 A (THE DISTILLERS COMPANY LIMITED) 8 July 1964
- D13: US-A-3 314 953 (VAZAKAS ARISTOTLE J ET AL) 18 April 1967
- D14: US-A-5 789 434 (KLUENDER ET AL) 4 August 1998 (1998-08-04)
- D15: SUZUKI, MAMORU ET AL: "Use of a new protecting group in an attempted

synthesis of cyclopropylidihydroxyphenylalanine" JOURNAL OF ORGANIC  
CHEMISTRY, 48(24), 4769-71 CODEN: JOCEAH; ISSN: 0022-3263, 1983,  
XP008044808

D16: SEDLAK M ET AL: "Preparation, <sup>1</sup>H and <sup>13</sup>C NMR spectra of substituted 2-  
benzoylaminocarboxamides" COLLECTION OF CZECHOSLOVAK  
CHEMICAL COMMUNICATIONS, ACADEMIC PRESS, LONDON, GB, vol.  
60, 1995, pages 150-160, XP002136842 ISSN: 0010-0765

D17: US-A-4 092 147 (ASHKAR ET AL) 30 May 1978 (1978-05-30)

D18: DE 33 32 633 A1 (LUITPOLD-WERK CHEMISCH PHARMAZEUTISCHE  
FABRIK GMBH & CO) 4 April 1985 (1985-04-04)

D19: WO 97/12625 A (CYTOVEN J.V; GREEN, LAWRENCE, R; BLASECKI,  
JOHN, W) 10 April 1997 (1997-04-10)

D20: US-A-4 291 048 (GOLD ET AL) 22 September 1981 (1981-09-22)

D21: WO 99/13873 A (EVLANKOVA, KLAVDIA STEPANOVNA) 25 March  
1999 (1999-03-25)

D22: US-A-5 783 605 (KUO ET AL) 21 July 1998 (1998-07-21)

D23: FOLKES L K ET AL: "Oxidative activation of indole-3-acetic acids to  
cytotoxic species- a potential new role for plant auxins in cancer therapy."  
BIOCHEMICAL PHARMACOLOGY. 15 JAN 2001, vol. 61, no. 2, 15 January  
2001 (2001-01-15), pages 129-136, XP008044869 ISSN: 0006-2952

D24: KARBOWNIK M ET AL: "Indole-3-propionic acid, a melatonin-related  
molecule, protects hepatic microsomal membranes from iron-induced  
oxidative damage: relevance to cancer reduction." JOURNAL OF  
CELLULAR BIOCHEMISTRY. 2001, vol. 81, no. 3, 2001, pages 507-513,  
XP008044870 ISSN: 0730-2312

D25: "THE MERCK INDEX" 2001, MERCK & CO., XP002322662

2 NOVELTY (Article 33(2) PCT)

2.1 The present application does not meet the criteria of Article 33(1) PCT, because the

subject-matter of claims 1-3,5,8,10-12,14-16,20 (as far as it relates to invention I) is not new in the sense of Article 33(2) PCT.

The document D1 discloses the antiangiogenic and antitumor activities of the thalidomide analogs of formula (V): CPS13,CPS16 and CPS20.

2.2 The present application does not meet the criteria of Article 33(1) PCT, because the subject-matter of claims 1,10,22 (as far as it relates to invention I) is not new in the sense of Article 33(2) PCT.

The document D2 discloses the antiangiogenic and antitumor activities of 3-amino-thalidomide and 2,6-diamino-thalidomide (compounds of formula (I)).

D3 and D4 disclose the antiangiogenic and antitumor activities of 3-amino-thalidomide (compound of formula (I)). D6 and D7 describe the antineoplastic activity of 4,5-dimethoxy-thalidomide (D6, table 1) and 4-methoxy-thalidomide (D7, table 2).

2.3 The present application does not meet the criteria of Article 33(1) PCT, because the subject-matter of claims 17,21 is not new in the sense of Article 33(2) PCT.

Document D5 discloses the azide derivative of thalidomide of formula V (see page 1281, scheme 2, compound 16).

2.4 The present application does not meet the criteria of Article 33(1) PCT, because the subject-matter of claims 1,10,22 (as far as it relates to invention II) is not new in the sense of Article 33(2) PCT.

Documents D9 and D10 disclose compounds of formula (I) wherein R40 is a substituted 5-membered, N-containing heterocyclic ring having marked anti-tumour actions (see the corresponding passages cited in the search report).

2.5 The present application does not meet the criteria of Article 33(1) PCT, because the subject-matter of claims 1,10,22 (as far as it relates to invention II) is not new in the sense of Article 33(2) PCT.

Document D11 discloses compounds of formula (I) wherein R40 is a substituted 4-membered, N-containing heterocyclic ring for treating various cancers (see the corresponding passages cited in the search report).

2.6 The present application does not meet the criteria of Article 33(1) PCT, because the

subject-matter of claim 22 (as far as it relates to invention II) is not new in the sense of Article 33(2) PCT.

Documents D12 and D13 disclose compounds of formula (I) wherein R40 is a substituted 5-membered, N-containing heterocyclic ring, and more particularly the compound 3-phthalimido-pyrrolidine-2,5-dione (D12).

2.7 The present application does not meet the criteria of Article 33(1) PCT, because the subject-matter of claims 1,10 (as far as it relates to invention III) and 23 is not new in the sense of Article 33(2) PCT.

Document D14 discloses compounds of formula IV, wherein R25 is methyl and R26 a substituted alkyl acid, for treating tumor metastasis (column 166, example 404; claims 1,8,10).

2.8 The present application does not meet the criteria of Article 33(1) PCT, because the subject-matter of claim 23 is not new in the sense of Article 33(2) PCT.

Documents D15,D16,D17 and D18 disclose compounds of formula IV (see the corresponding passages cited in the search report).

2.9 The present application does not meet the criteria of Article 33(1) PCT, because the subject-matter of claims 1,10 (as far as it relates to invention IV) and 24 is not new in the sense of Article 33(2) PCT.

The document D19 discloses the antiangiogenic and antitumor activities of a compound of formula IX, the dipeptide L-Glu-L-Trp (R33 is an amino acid moiety, i.e. any moiety having a carboxylic and an amino group).

2.10 The present application does not meet the criteria of Article 33(1) PCT, because the subject-matter of claims 1,10 (as far as it relates to invention IV) and 24 is not new in the sense of Article 33(2) PCT.

Documents D20 and D21 disclose the use of a compound of formula IX (i.e. L-tryptophan) for treating tumors.

2.11 The present application does not meet the criteria of Article 33(1) PCT, because the subject-matter of claims 1,10 (as far as it relates to invention IV) and 24 is not new in the sense of Article 33(2) PCT.



Documents D22, D23 and D24 disclose the compounds of formula IX, indole-3-acetic acid (D22,D23) and indole-3-propionic acid (D24), for use in cancer therapy.

2.11 The present application does not meet the criteria of Article 33(1) PCT, because the subject-matter of claim 24 is not new in the sense of Article 33(2) PCT.

Document D25 discloses compounds of formula IX: indole-3-acetic acid, indole-3-propionic acid and tryptophan.

The attention of the Applicant is drawn to the fact, that the mere explanation of an effect obtained when using a compound in a known process, even if the explanation relates to a pharmaceutical effect which was not known for that compound, cannot confer novelty to said process. In the present case, the newly discovered technical effect of inhibiting angiogenesis does not confer novelty on the claims directed to the use of a known thalidomide derivative for a known purpose (treatment of tumors or cancer). No novelty exists, if the claim is directed to the use of a known compound for a known purpose, even if a newly discovered technical effect (angiogenesis inhibition) underlying said known use is indicated in that claim.

### **3 INVENTIVE STEP (Article 33(3) PCT)**

3.1 Should the Applicant have overcome the objections of lack of novelty raised above, an inventive step could not be acknowledged over D1 to D4 (invention I), D9,D10,D11(invention II),D14 (invention III), D19 to D24 (invention IV) as the present subject-matter of claims 1-3,5,8,10-12,14-17,20-24, as far as novel, appears to be an obvious alternative over said documents (Article 33(3) PCT).

The antiangiogenic activity of compounds of formula I, V and IX has been clearly and unambiguously disclosed in D1,D2,D3,D4 and D19.

The antitumor and anticancer activities of compounds of formulae I (wherein R40 is a 4-or 5-membered, N-containing heterocyclic ring) and IV has been clearly and unambiguously disclosed in D9,D10,D11 and D14.

3.2 The same applies to the subject-matter of the dependent claims 4,6,7,9,13,18,19 which apparently does not contain any technical features which could be regarded as inventive per se. The skilled person would have expected for the compounds recited in said claims the same activity in the inhibition of angiogenesis as the closely related thalidomide analogs of D1 to D4.

#### **4 INDUSTRIAL APPLICABILITY (Article 33(4) PCT)**

4.1 There are no doubts about industrial applicability (Art.33(4) PCT) for the subject-matter of claims 14-24.

4.2 However, for the assessment of the present claims 1-13 on the question whether they are industrially applicable, no unified criteria exist in the PCT Contracting States. The patentability can also be dependent upon the formulation of the claims. The EPO, for example, does not recognize as industrially applicable the subject-matter of claims to the use of a compound in medical treatment, but may allow, however, claims to a known compound for first use in medical treatment and the use of such a compound for the manufacture of a medicament for a new medical treatment.

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☒ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☒ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**